UiO : Faculty of Mathematics and Natural Sciences
University of Oslo

Visualizing Subduction Using Statistical Modelling Applied to the Cascadia Slab

Evmorfia Andritsopoulou

Master's Thesis Spring 2015

Visualizing Subduction Using Statistical Modelling Applied to the Cascadia Slab

Evmorfia Andritsopoulou

24th May 2015

Abstract

The intention of this thesis is to give a better understanding of the way that subduction zones evolve, to examine the way that different subduction parameters affect each other and finally to demonstrate how visualization can be used as a tool to provide deeper insight into such zones.

The first part, describes the characteristics and the theory behind the formation and evolution of the areas that the phenomenon of subduction takes place. It can be especially useful to readers who do not have extensive previous knowledge on this subject.

The second part, uses 20 measurable parameters of subduction zones to develop statistical models in order to reveal correlations and tendencies within geological observations around zones of subduction. This models are created using multiple linear regression with the help of the R statistical software environment.

The third and final part, deals with the visualization of geological phenomena by using software that computer science has developed for geoscience. For this purpose, bathymetrical reconstruction of the Cascadia slab is performed and visualized, using the previously acquired models and the 4DPlates plate reconstruction software.

Contents

List of Figures

List of Tables

Preface

This thesis is the final work of my studies in Computational Science at the University of Oslo for which I collaborated with the department of Computational Geoscience of Simula Research Laboratory and with Kalkulo AS. It consists of the documentation and the results of my research during the last year and a half of my studies.

My interest in computer science and mathematics started at an early age in my life and my later studies naturally followed this path. For my thesis I applied these two disciplines on the basis of the geological phenomenon of subduction as geoscience was always a field that fascinated me. It was a pure learning experience in all possible aspects and is the reason why I now feel satisfied with myself for my work.

Nothing would be possible without my supervisors Stuart Clark, Karsten Trulsen and Are Magnus Bruaset. My main supervisor Stuart was the one to motivate me, support me by all means and advise me whenever I needed it. He was also the one who proposed the initial idea about the focus of this work. Karsten also helped me greatly with his always quick and insightful responses as well as his positive attitude. Simula and Kalkulo AS were also important contributors during these years providing me with all the needed material and scientific support .

I would also like to thank George and my friends from Greece and Simula for accepting my madness and emotional ups and downs kindly and giving me back actual support and advice.

Finally, my family deserves a special and big thank you as both my parents -Christos & Dimitra- and my two brothers -George & Theodore- supported me from the beginning of my studies practically and emotionally even when I did not ask for it.

Part I

Introduction

Chapter 1

Seismic Waves

Even from the first years of geological exploration, scientists were concerned and curious about the inner structure of the earth. The English scientist Isaac Newton was the first to deal with this subject. Through his studies of planets and the force of gravity he concluded that the average density of the Earth is more than twice the density of the rocks near the surface and that made him realize that the interior of the Earth is composed of much denser material than the surface rocks. Trying to establish how this mass is distributed and thus get a picture of the inner structure of the Earth took many years due to technological boundaries and shaped the science of seismics as we know it.

After every earthquake, waves of energy that travel through the Earth are generated. These waves are called seismic waves and due to their nature to travel at different speeds in different materials, help us understand the internal structure of the earth. The two main types of such seismic body waves are the P-waves (primary or compressional waves) and S-waves (secondary or shear waves). P-waves travel through all kind of materials including gases, liquids and solids and travel relatively fast at speeds between 1 to 14 km/s depending on the rock type. The motion which is produced from such a wave is an altering compression and expansion of the material. On the other hand S-waves travel slower at speeds between 1 to 8 km/s within the Earth and are incapable of travelling through liquids. These waves' movement is perpendicular to the direction that the wave is travelling. [34]

The change of speed and direction of the waves when they pass through to another material and the difference of speed of travel between P-waves and S-waves, provide scientists a view of the inner structure of the Earth. The fact that both P- and S-waves are detected in seismometers has determined that the mantle of earth is solid in contrast with the outer core which is "molten" or liquid as the S-waves that travel into it cannot be detected on the other side of the Earth. If we measure the time it takes for a seismic wave to travel through Earth, we can easily determine the velocity structure of the Earth. All these measurements are of course taken with the help of seismographs, recording every earthquake. [57] [3] [58]

Chapter 2

Seismic Tomography and Imaging

Seismic tomography is a technique that images the Earth's interior structure using seismic waves generated by earthquakes and explosions. Seismic tomography has a medical analogue which is CAT (computer – aided tomography) scanning as they both combine information from large numbers of intersecting waves in order to build three dimensional images of the medium that the rays have travelled through.

Almost every image of seismic tomography is based on the spatial distribution of the velocity of seismic waves which is determined by using travel time data. This data is acquired from an array of seismographic stations placed all over the Earth's surface [3]. In order to map the 3-D distribution of the P and S-waves' velocities as well as the locations of discontinuities that happen at interfaces between different layers, we have to analyse the arrival times of waves. The variation in the velocity of the waves inside the Earth is mostly related to the temperature and compositional variations that occur. In lesser extent, wave propagation velocity depends on a small-scale property of the medium which is the orientation of crystals in it.

- **Temperature Variations**

  Colder materials generally tend to be harder and more resistant to compression compared to hot ones. That is why seismic waves travel through cold areas of the Earth's interior more rapidly. On the other hand warmer materials have softer consistency and as a result the waves travel more slowly.

- **Compositional Variations**

  In order to explain the way that the seismic tomography images the Earth's interior we have to use the principles of wave propagation through different media. From Snell's law, we know that when a wave crosses an interface between two isotropic media, the wave changes its direction according to the following formula and this applies to both P- and S- waves.

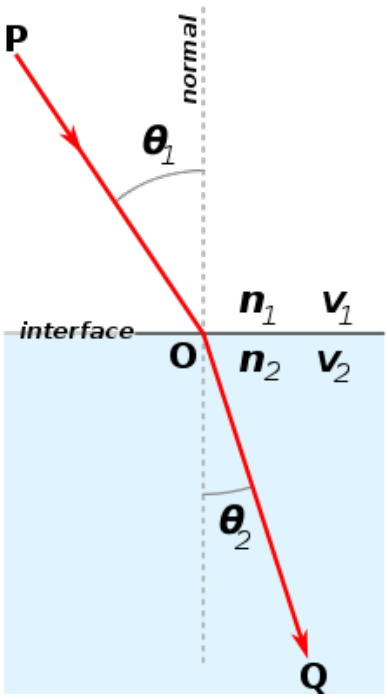


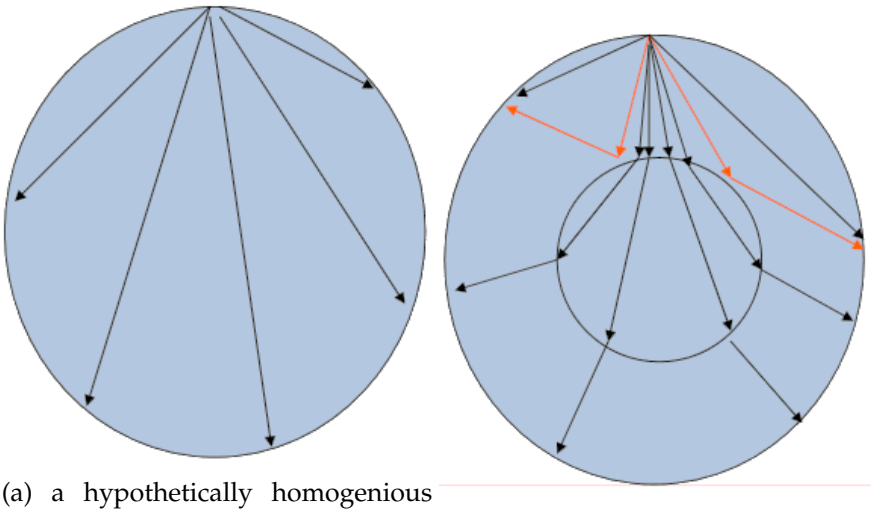
Figure 2.1: Snell's Law [54]

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{V_1}{V_2} = \frac{n_2}{n_1}$$

Where V_1, V_2 the velocity of light in the respective medium and n_1, n_2 the refractive index (which is unit-less) of the respective medium. Graphically Snell's law is depicted in 2.1.

If the Earth was homogeneously composed throughout its spherical body the seismic waves would travel in straight lines and we would only deal with geometrical relation between the P and S-waves' travel times and the epicentral distance (see figure 2.2a). In reality the inside structure of the Earth is divided in layers and that complicates the relation between the waves' travel times and the epicentral distance following Snell's law (see figure 2.2b).

All the previous make obvious that the velocity of seismic waves contains indirect information about the Earth's internal flow. The way to extract these information though is anything but easy. A single ray's velocity computed at the time of arrival at a seismic station is only an average velocity over the entire path that the ray travelled and does not reveal the areas where the wave has been decelerated or accelerated. In addition, the average velocity is normally calculated over great distances because of two reasons. Firstly, there are large expanses on the Earth without any seismic station, especially because of the oceans and due to economic and political reasons. Secondly, earthquakes mostly occur around plate boundaries and they are impossible to be predicted.



(a) a hypothetically homogenious Earth

(b) the layered Earth

Figure 2.2: The way that seismic waves travel through [35]

Nevertheless, the vast amount of data in the seismic database, gives the opportunity to scientists to construct detailed images of the Earth's internal seismic-velocity structure applying tomography. [16], [1], [34]

Using all the knowledge gained through the seismic waves and laboratory experiments, scientists discovered that the Earth's interior is anything but homogeneous and that it is made up of distinct layers like shown in the figure 2.3. The way we can define the Earth's interior structure though is dual. Firstly we can do it by using mechanical properties like rheology and secondly by using chemical properties. Chemically, the principle layers beginning at the centre of the earth are the core, the mantle and the crust. Mechanically the layers are the core, the mesosphere, the asthenosphere and the lithosphere [11], [24]. The connection between the chemically divided layers and the mechanically divided ones is depicted in figure 2.3 and explained in details below.

2.1 Core

The core of the earth is approximately 3500km thick and is considered to be mainly composed of nickel and iron alloy. This assumption is based on calculations according to its density and on the fact that many meteorites which are considered to be portions of the inner part of a planetary body are iron-nickel alloys. The Earth's core contains radioactive materials which break down into more stable substances and release heat. That makes the core the Earth's source of internal heat.

The core is divided in two different parts, the outer and the inner core. The outer core is electrically conducting liquid as the extreme temperatures are adequate to melt the iron-nickel alloy. The outer core is the only liquid layer of the earth, it is about 2300 km thick and is located 2890 to 5150 km below the surface of the earth On the contrast the inner core is solid even

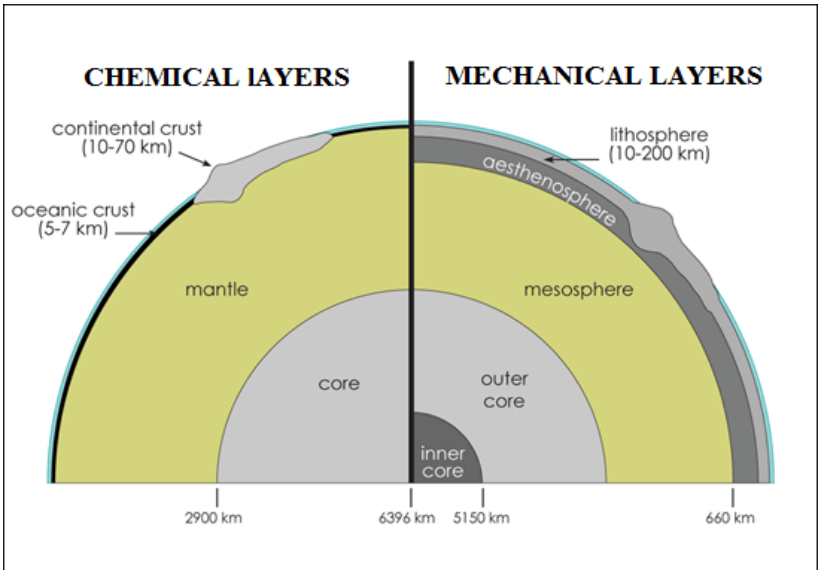


Figure 2.3: Chemical and Mechanical layers of Earth's interior [62]

though the temperature is much higher than the one of the outer core. The reason for that is the tremendous pressure of the overlying rocks which is strong enough to crowd the atoms tightly and form a solid state. The inner core's consistency is mostly of iron and nickel, its diameter is about 1200km and it is located 5150 to 6378 km below the surface of the Earth. [31], [22], [25]

2.2 Mantle

The Earth's mantle is approximately 3000 km thick, it is thought to be consisted of mainly olivine-rich rock and has different temperatures at different depths. In general terms the temperature increases with depth and the highest ones occur where the mantle material meets the heat-producing core. This correlated increase of temperature and depth is known as geothermal gradient which causes different rock behaviours and these behaviours are used to distinguish the mantle in two different parts, the upper mantle and the lower one.

The upper mantle consists of rocks that are cool and brittle which makes them break under stress and produce earthquakes. However the rocks in the lower mantle are hot and soft –semisolid, not molten, so they can flow instead of breaking when they are subjected to high forces. [31], [22], [26]

2.3 Crust

The crust is the Earth's outermost and thinnest layer and its consistency is hard and rigid. We distinguish the crust in two different types, the oceanic crust that underlies the ocean basins with only 5 to 7 km thickness and the continental crust which underlies the continents and has 10 to 70 km

thickness. These two different crust types are composed of different rock types. The thick continental crust is primarily composed of granite and its low density allows it to float on the much higher density mantle that is located below. The thinner oceanic crust is primarily composed of basalt. [31], [22], [26]

2.4 Lithosphere

The lithosphere is composed of the crust and the upper mantle, it constitutes the harder and more rigid outer layer of Earth. Like the crust, we can distinguish lithosphere in oceanic and continental one. Oceanic lithosphere exists in the ocean basins and is typically about 50 to 140 km thick. Continental lithosphere underlies continents and its thickness ranges from 40 km to approximately 280km. When it comes to continental lithosphere, the upper 30 to 50 km are the crust. [38]

The lithosphere is broken up into giant rigid pieces which are called tectonic plates and move slowly each year as they slide on top of a part of the mantle that is called asthenosphere.

2.5 Asthenosphere

The asthenosphere lies directly below the lithosphere and is a portion of the upper mantle. It lies below the lithosphere at depths between 80 and 200 km below the surface. Its thickness depends mostly on the temperature but in some regions, asthenosphere can be 700 km thick. It is a malleable semi-liquid zone and a small percentage of melt makes the seismic waves travel relatively slowly through this layer compared to the overlying lithosphere. The reasons for this ductile state are the temperature and pressure conditions that turn the rock into a semi-fluid which moves forming currents. [26]

2.6 Mesosphere

The mesosphere is the part of the mantle below the asthenosphere but above the outer core. In simple terms it can be described as: Mesosphere = (upper mantle + lower mantle) – (lithosphere + asthenosphere)

Chapter 3

Plate Tectonics

3.1 Early Days

In the early 1900s, a German scientist Alfred Wegener (1880-1930) noticed that most of the continents seem to fit together like a puzzle especially when comparing the continental shelves instead of the coastlines. Because of this observation he proposed the idea that the continents were once forming one single protocontinent which he named Pangaea and over time they split and moved apart into their current positions. Wegener's hypothesis also explained the way that the mountains were formed (orogenesis). He explained that as the continents were moving, their leading edges were encountering enormous resistance which caused compression, upwards fold and finally mountain formation. The prevailing theory until that time was the "contraction theory" which stated that the planet was once a molten ball and during the process of cooling down the surface cracked and folded up on itself forming the mountains. This theory though was not explaining why all the mountains did not have approximately the same age. Finally Wegener proposed that the mechanism that forced the continental break up and drift was a centrifugal force caused by the rotation of the earth.

In 1929, Arthur Holmes was the one who came up with the idea that the mantle undergoes thermal convection. This phenomenon occurs while we heat a substance and its density decreases. This makes the substance rise to the surface until it is cooled down and sinks again. This current was responsible, according to Holmes, for breaking up the continents and moving them apart. [67], [25]

3.2 Modern Days

The modern plate tectonics theory was widely accepted at 1960s and states that the Earth's outer part, the lithosphere, is divided into large slabs which are called plates. The lithosphere can be divided in eight major plates and many minor ones. These plates underlie the oceans and the continents and are slowly but constantly moving (typically from 10 to 150 mm per

year). This movement can explain many geological events that occur, like earthquakes and volcanoes.

The location that the plates meet is called plate boundary and the relative motion of the plates in that area determines the type of boundary; convergent, divergent or transform [44]. A convergent boundary occurs where two plates are moving towards each other. In a divergent boundary the two plates are moving apart from each other and finally at a transform boundary the two plates slide past each other.

3.2.1 Convergent Boundary

A convergent boundary can also be called a destructive plate boundary. This is a highly deforming region where at least two tectonic plates move towards one another and collide. Due to this collision one of the tectonic plates is "forced to subduct under the other" and this is how a subduction zone is formed (see figure 3.1a). The result of the pressure and the frictions around the subduction zones are earthquakes and volcano forming. The plate that subducts in these zones is normally a plate with oceanic crust and moves beneath a plate with oceanic or continental crust. More information about subduction theory is given in chapter 4 In the cases that the two colliding plates are both made of continental crust, it is not referred to as a subduction zone but as a continental collision (see figure 3.1b). During these collisions large mountain ranges are formed, a good example of this are the Himalayas. [44], [69]

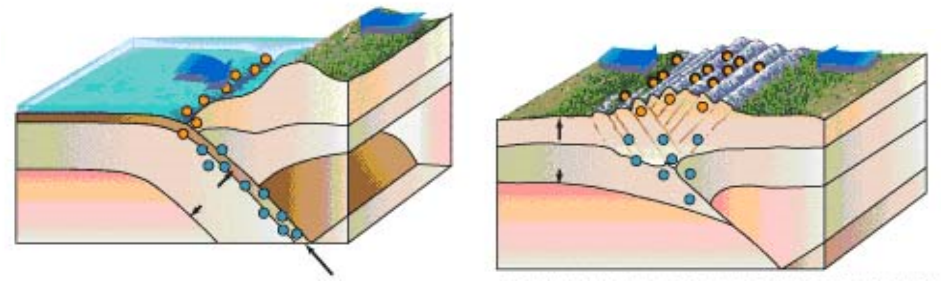
3.2.2 Divergent Boundary

A divergent boundary can also be called constructive boundary or extensional boundary and occurs between tectonic plates that are moving away from each other (see figure 3.1c). Within continents that diverge, rifts are initially formed which later become rift valleys. The most actively diverging plate boundaries are the ones between oceanic plates and form mid-oceanic ridges.

In divergent zones, the motion that pulls away the two plates creates a space between them. This space reveals the deep mantle rock of the asthenosphere, the molten magma. As this magma rises to the top, it freezes onto the trailing edges of the diverging plates, filling the gap that was created and expanding the plates. In that way new lithosphere is created with hot material and over millions of years it cools down. While it cools down it shrinks more and more and that is why fresh sea floor always stands higher than the older lithosphere and mid-ocean ridges take the form of long and wide swells. Divergence that happens between continental plates is the reason why new oceans are born [44], [60].

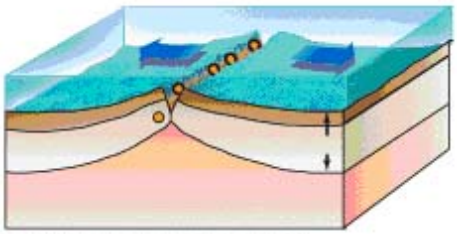
3.2.3 Transform Boundary

Transform boundaries are also called transform faults or conservative plate boundaries and are places where the plates move sideways past each

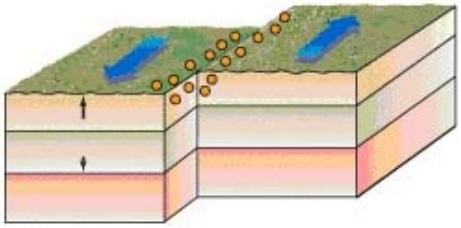


(a) Subduction zone

(b) Continental Collision



(c) Divergent



(d) Transform Fault

Figure 3.1: Different types of plate boundaries [41]

other (see figure 3.1d). At these boundaries lithosphere is neither created nor destroyed like at divergent or convergent boundaries respectively. California's San Andreas fault is one of the most well-known transform boundary. Transform boundaries end abruptly and are connected on both ends with either other faults or ridges or subduction zones. Transform faults help the strain relief caused by compression, extension or lateral stress in the rock layers, by transporting it between ridges or subduction zones [26].

Chapter 4

Subduction Theory

Subduction is a scientific word from the Latin language and means "carried under". Subductions happen in convergent boundaries where one plate meets another and one sinks under the other, into the mantle (see figure 4.1). The regions where this phenomenon occurs are known as subduction zones. Typically, the rate that subduction occurs is some centimetres per year and the average rate of convergence is approximately between 2 to 8 cm per year.

When two continents meet, there occurs no subduction as the continents are made of rocks that are too buoyant to sink more than about 100 km deep. When oceanic lithosphere meets continental one, the continent always "wins" and it is the oceanic plate that subducts. The last possibility is that an oceanic plate meets another oceanic plate. Here it is the younger plate that "wins" and the reason for that is the density. When the oceanic lithosphere is formed at the mid-oceanic ridges, it is thin and hot but as it moves away from the ridge it cools down and becomes thicker because more rock hardens underneath it. As all the rocks do, while lithosphere cools, it shrinks more and that is why it sits lower. As the years pass, the oceanic lithosphere becomes denser and colder than the hot asthenosphere that lies beneath. For this reason when two oceanic plates meet, it is the younger and higher density plate that has the advantage.

Once a plate begins to subduct, it is the gravity that takes over and when it is already descending, it is usually called "slab". The sinking plate can form an angle of approximately 25 to 45 degrees to the Earth's surface. It is observed that when the sea floor is very old, the slab sinks almost straight down. On the contrary, when the plates are quite young, the slab sinks at a shallow angle.

When it comes to gravitational "slab pull", subduction is considered to be the strongest force driving plate tectonics and without it plate tectonics would not occur. The cause for the plate sinking and "slab pull" is the temperature difference between the mantle asthenosphere and the oceanic lithosphere that subducts as the oceanic lithosphere is colder and on average denser. The high pressure that occurs at a depth of approximately 80 to 120 km, turns the basalt that the oceanic lithosphere is made of, into a denser metamorphic rock called eclogite. As the density is increased, it

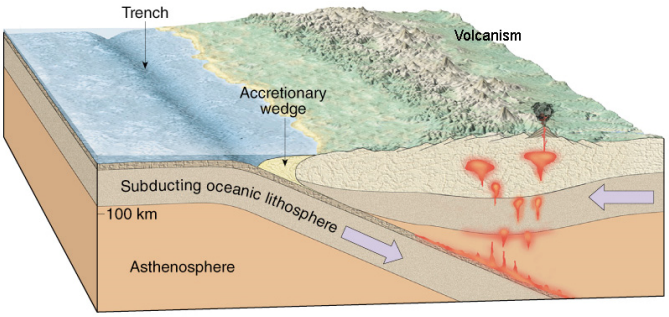


Figure 4.1: Subduction Zone [59]

provides the slab with an additional negative buoyancy and this makes the subducted slab much more eager to descend.

At the point that a slab starts to bend downwards, a deep-sea trench is formed. Trenches tend to capture a lot of sediment from the nearby land masses, much of which is carried down with the subducted plate. There is also another possibility for the sediment which is captured. While the one tectonic plate subducts, it floats on the asthenosphere and so it pushes against the plate which lies on the top. This can cause the scrapping off of sediments from the slab by the top plate. These sediments form a mass of material called accretionary wedge which attaches to the end of the upper plate, forming a wedge like soil in front of a plough. This is demonstrated in figure 4.1

Subduction zones are also areas with high rates of volcanism, earthquakes and orogenesis. Above subduction zones, exist long chains of volcanoes which are called volcanic arcs. These volcanoes tend to be extremely explosive and produce dangerous eruptions because of the water content from the slab and the sediments. In addition, arcs are associated with precious metals like gold, silver and copper, which are believed to be carried out by the water and can be found in rocks called ore in rock terminology. Orogenesis or else mountain-building, takes place when large pieces of material on the subduction plate are pressed into the plate which over-rides or when the over-riding plate contracts sub-horizontally The interactions between the slab and the mantle, the volcanoes and the mountain-building are the reasons why the areas around subduction zones are subject to many earthquakes. [8], [28], [34]

4.1 Origin Theory

It is true that the initiation of a subduction is most probably the most poorly-understood phenomenon of plate tectonic theory. While all phenomena like opening and closure of ocean basins indicate that the initiation of a subduction is common, theoretical and mathematical models show that it is quite difficult for a subduction to be initiated. Nevertheless, as many scientists say, without subduction zones there would not be such

a thing like plate tectonics.

The question of how subduction is initiated has been a matter of considerable debate. There are two main lines of thought that try to answer this question. The first explanation, which is also the most commonly used, is that as the oceanic lithosphere ages, it also cools down and as an effect its density increases causing an instability and the spontaneous sinking of the plate. The problem with this explanation is that statistical modelling has proved that at a fracture zone it is highly unlikely that the entire lithosphere will start sinking spontaneously. Scientists claim that without the existence of convergence, fault rheology or geometry alone are not enough to initiate a self-sustaining subduction.

The second explanation dictates the need of both moderate convergence and compressive stresses applied from an external source to take place for a new subduction zone to be formed. A mechanism that is the best candidate to generate the previous forces is a collision. The stresses produced from the collision would be transferred forcing the initiation of a subduction elsewhere. Nevertheless, the results from recent collisions show something different. About 35 to 50 million years ago the closure of Tethys Ocean included the collision of India and Africa with Eurasia. The previous theory dictates that large-scale collisional stress transfer should occur, resulting in the initiation of a subduction somewhere within the Indian and the African plates. However, after 50 million years, there is no creation of a new subduction zone and the only geophysical event that happened and changed the geometry of the area is the formation of the Alpine-Himalayan chain.

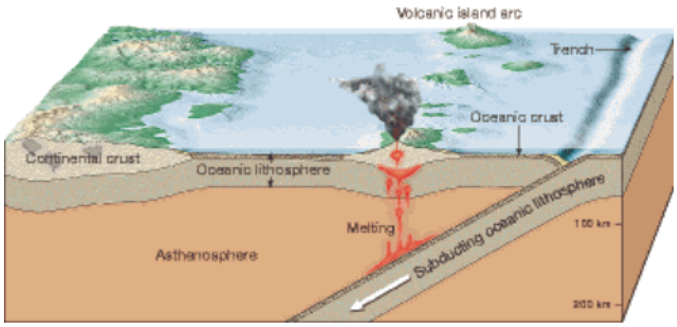
It is a common assumption that subduction is continuously occurring on Earth but it is also a fact that supercontinents are constantly created and oceans close. The only new subductions that have occurred within the past 80 million years are firstly and most importantly the 600 km long Scotia Arc and an intra-oceanic subduction which is located in the Pacific basin.

As mentioned before, scientists use models to explain the way that subductions occur and develop but they have to face some practical issues. These models, in order to achieve plate-like surface motion utilise boundary conditions that are determined dynamically. It is though difficult to imitate and reproduce the terrestrial convective energy. It is impossible to model the dynamics that are responsible for the creation of the plates and they only provide insight to the history of the dynamics of both the plates and the mantle since the existence of the plates is assumed.

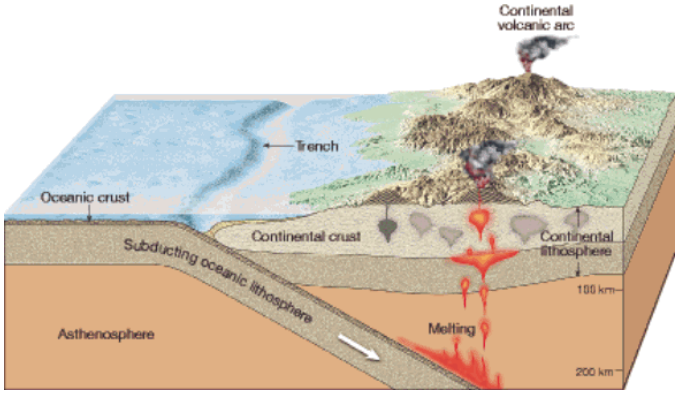
The result of all the previous is that placing the plate tectonics theory in the world of physics, is anything but simple and some believe that it is even impossible. All things considered, it is impressive how little progress there has been in modelling in this field. [17], [19], [36], [53]

4.2 Physical Effects

For geologists, identifying a subduction zone is quite an easy process. There are four main indicators that indicate an area is a subduction zone.



(a) Volcanic Island Arc



(b) Continental Volcanic arc

Figure 4.2: Volcanic Arcs [9]

These are: the volcanic activity, seismicity and tsunamis around the area, the mountain formation and deep sea trenches.

4.2.1 Volcanic Activity

An important tectonic setting where many volcanoes occur is around subduction zones and they represent about 10 to 20 percent of the volcanism on Earth. The oceanic crust that sinks into the mantle in a subduction zone contains large amounts of surface water, carbon dioxide and volatile elements which are contained in hydrated minerals within the basalt that the sea floor is made of. While the slab descends into the mantle, it encounters progressively great pressure and temperature levels and these make the slab release water into the overlying mantle wedge that the slab forms with the upper plate. The addition of fluids in the slab lowers the melting temperature of the mantle (similarly with the melting temperature of the ice when salt is added) and this results in the melting of the slab and facilitates the magma generation.

The composition of magma that this mechanism produces, is a variation of basalt and andesite. The fact that the magma is lighter and less dense than the rocks of the mantle, makes it rise upwards, above the subduction zones and form a linear belt of volcanoes which lie parallel to the oceanic trench. If the slab subducts under oceanic crust the volcanic arc which

forms, is called volcanic island arc (see figure 4.2a). and an example of this is the Aleutian Island Chain. In the case that the slab subducts under continental crust the volcanic belt which forms is called continental volcanic arc (see figure 4.2b) and examples of such belts are the Cascades volcanic arc of the U.S. Pacific northwest and the Andes volcanic arc of South America. Volcanoes around the subduction zones are typically very explosive. As the magma travels up to the surface, it undergoes a variety of processes like cooling and partially crystallizing. At different depths the magma pools and cools down and that causes its partial crystallization. Because of the formation of different crystals, the remaining fluid magma, changes its original hot basaltic chemical composition (silica poor, iron and magnesium rich), to more silica-rich compositions (andesitic, dacitic or even rhyolitic). The silica-rich composition makes the magma very viscous and in the same time, as explained before, it contains many fluids (water, carbon and sulphur dioxide). At the surface, these fluids form bubbles which are unable to escape -because of the viscosity- in any other way but explosive fragmentation.

Volcanoes in subduction zones are also called "gray volcanoes" in contrast with the "red volcanoes", as their eruptions usually produce gray ash plumes . [52], [40], [50]

4.2.2 Earthquakes

The earthquakes that occur at subduction zones and are caused because of the overriding plate slipping over the top of the subducting slab are called mega-thrust earthquakes. It is a fact that nine out of ten earthquakes that occurred in the last century were subduction zone events. Within these, is included the largest ever recorded earthquake in the 1960s which is the Great Chilean Earthquake with M [1] 9.5.

In subduction zones, at the point that the two plates converge, stress is generated as large sections of the two plates become locked together and this prevents the plates from slipping smoothly at their boundaries. This builds up stress within the rocks and large amounts of energy are stored until they break, creating great earthquakes of magnitude often around M9.0 or higher. It is basically the energy that is released during the sinking of the slab that is mostly used to move and deform the tectonic plates. Part of this energy is used to overcome resistance at the subduction zone plate boundary and part of this energy is released in the form of subduction zone thrust earthquakes

The zone of seismicity that is formed on the down-going slab in a subduction zone is called Wadati-Benioff zone and produces numerous earthquakes which can be as deep as about 670 km because of the shortening and extension of the plate that is pulled into the mantle as shown in figure 4.3. The Wadati-Benioff zone which lies on the angle of dip of the slab that subducts is controlled by the negative buoyancy of the

[1] M or M_w: Moment magnitude scale, successor to the Richter scale. It does not only measure the amplitudes of the recorded waves but also the energy released, taking into account what physically occurs during an earthquake.

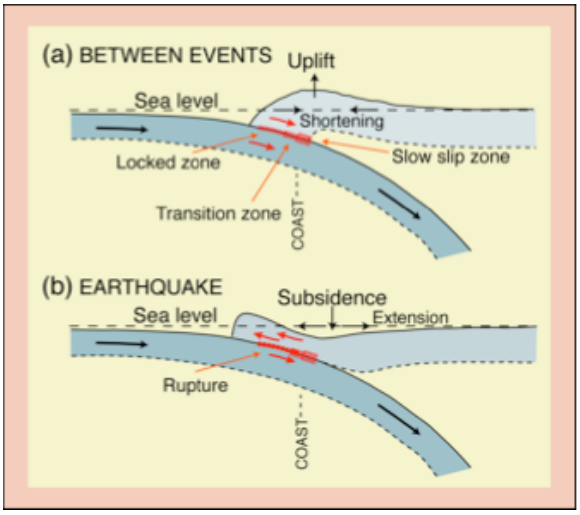


Figure 4.3: Shortening and extension of the slab generates earthquakes [21]

slab and the force that is created by the flow of the asthenosphere. When lithosphere is younger it also has higher temperature and is more buoyant and this results in shallow-dipping Benioff zones. On the contrary, older lithosphere is colder and denser which causes steeper dips. The zones have dips that can typically range from 40 to 60 degrees.

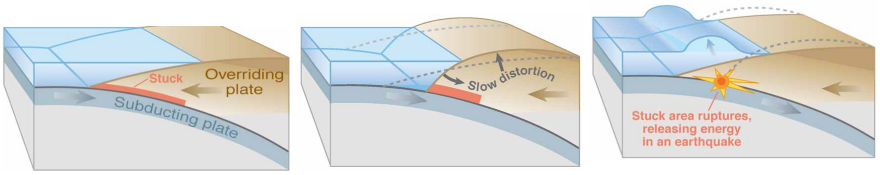
Scientists have noticed that different subduction zones show differences in seismic behaviour and try to point out what are the factors that affect the magnitude of the earthquakes that take place around the subduction zones. There are subduction zones which produce mega-thrust earthquakes, like Chile, Alaska, Sumatra-Andaman and Japan, while others like Scotia, produce relatively smaller earthquakes.

Approximately 80% of all the earthquakes that have been recorded have occurred in the Pacific Ocean basin as there are many subductions that take place there. The three largest recorded subduction zone thrust earthquakes are the 1960 M9.5 Central Chile earthquake mentioned earlier, the 1964 M9.2 Alaska earthquake and the 2004 M9.1-9.3 Sumatra-Andaman earthquake. Finally, scientists try to calculate the earthquake supercycles [2] of different subduction zones, so that they can be able to predict future seismic events. For example along the Cascadia subduction zone there has not been any great instrumentally recorded (with magnitude M>= 8) subduction zone earthquake. There is evidence though that in the last 7500 years 13 seismic events have occurred which give average repeat times little less than 600 years. [26], [48], [33]

4.2.3 Tsunamis

For an earthquake to generate a tsunami we need two main factors; water and vertical motion. If an earthquake happens in an area that is far away from a body of water there is no disturbance of water and so no tsunami

[2] A sequence of massive earthquakes in a region that happen in a chain



(a) Subduction zone

(b) The overriding plate is squeezed and its leading edge is dragged down. The area above this curves upwards. This movement accumulates great stress between the plates.

(c) A subduction zone earthquake happens, and the leading edge of the overriding plate breaks free releasing all the accumulated stress in the form of energy. This energy raises the sea floor generating a tsunami.

Figure 4.4: Tsunami generation [13]

is expected to be generated. As for the second aspect, the vertical motion, it is only a convergent boundary that produces vertical motion capable to generate a tsunami. Earthquakes at transform boundaries involve nearly no vertical motion while divergent boundaries have some but not enough vertical motion. It is only convergent boundaries that always involve strong enough vertical motion and are capable of generating a tsunami like it is graphically explained in figure 4.4.

During a subduction, the overriding plate accumulates energy while it is locked with the subducting slab. At the moment that this energy exceeds the frictional forces between the two plates, the overriding plate snaps back into an unrestrained position. This sudden release gives enormous thrust to the overlying water and this generates a tsunami. At the same time the areas on the overriding plate are suddenly lowered.

The speed at which a tsunami travels is enormous. The tsunami wave that was produced by the subduction zone earthquake along the coast of Chile in 1960, reached Hawaii, after travelling across the Pacific Ocean, in only about 15 hours and Japan in less than 24 hours. [2]

4.2.4 Orogenesis

Orogenesis or orogeny derives from two Greek words which mean formation of mountains. The reasons that cause a mountain formation are many but plate tectonics is one of the most outstanding. At destructive margins where two plates collide, rocks are folded and lifted and so they form chains of fold mountains. It is therefore dynamic forces that thrust the land upward. Of course not all the mountains are a result of these forces but in this paragraph we will focus in mountains that result from a plate tectonic process.

There are two energy sources that initiate mountain building; heat and mechanical energy. Firstly we have the mountains that their formation

is a result of the heat from the inside of the Earth which warms up the lithosphere causing it to expand, lift and bend the surface upwards. Subduction volcanoes are usually associated with this form of energy, they are formed along subduction zones and they are called cordilleran mountains. Secondly, we have the mechanical mountain building that is composed of tension and compression. The differences in elevation or relief, are created when different parts of the earth move relative to each other - falling or rising vertically- , or when one part of the earth is being shoved -thrust- horizontally over another.

Subduction zones can be found anywhere in an ocean basin, can face any direction and more than one subduction can be active in the same basin at the same time. That may result in complex sequences of mountain buildings which are considered to be usual. Mountain building in convergent zones is a compressive mechanism and either one or both mechanically and heat driven mechanisms can occur for the formation of a mountain. The heat driven ones though are initially compression forces that ultimately generate heat.

There are two kinds of subduction orogeny, those which occur within two oceanic plates and form mountains within the ocean basins - island arc type- and those which occur when oceanic plate subducts under continental one and the mountains are formed on the land - cordilleran type-. In island arc type orogeny, the uplift is mostly heat driven as magma rises from the mantle. The most common mountains in island arcs are volcanic ones. Examples of such orogeny are the islands of Japan and the Aleutian islands. Cordilleran mountain building is also heat driven as the heat swells the continent upwards and then volcanoes can build even higher on top of that. Examples of such orogeny are the Andes Mountains or the Cascades. In both orogenic types, there are many processes that occur and they generate a wide diversity of rocks and structures.

Finally except from subduction orogeny we can distinguish the collision orogeny which can be divided in continent-island arc type and continent-continent type. In both cases the ocean basin descends under the continental plate until it completely disappears. These ocean basins are called remnant ocean basins (ROB) and the two land blocks on either side of the remnant ocean have no other choice but to collide. From this collision it is possible that we get a collision orogeny and wide variety of rocks and structures can be produced.

It is worth mentioning, that the longest and highest mountains on Earth are volcanic ones. Mauna Kea in Hawaii rises 4205 m above the sea level and is 6000 m deep. That makes it a total height of 10205 m compared to Everest which is 8850 m high. In Europe the highest volcanic mountain is Mt Elbrus in Russia with 5633 m height. [37], [45], [14]

4.2.5 Trenches

Trenches are distinctive morphological features of convergent plates and are also called submarine valleys. They are hemispheric-scale long but narrow and make up the deepest parts of the ocean floor.

Trenches are steep, V-shaped topographic depressions of the sea floor, that are formed at the position in which the subducting slab is bent and begins to descend under the other overlying plate as shown in figure 4.5. The speed in which the oceanic lithosphere disappears into trenches is about 3 km^2/yr at a global rate. Typically, trenches are parallel to a volcanic arc and the distance between them is about 200 km while they extend 3 to 4 km below the level of the oceanic floor that surrounds the trench.

Most of the trenches can be found in the Pacific Ocean, as it contains the most kilometres of convergent plate margins. The deepest trench of them all is also located in the western Pacific Ocean , to the east of the Mariana Islands and is called Mariana Trench. Its length is about 2550 km but it is only about 69 km wide. The spectacular characteristic though is the deepest known part of this trench - and of the ocean as well- which is called Challenger Deep and is around 10994 m deep. For comparison, the world's tallest mountain, Mount Everest, has a height of 8850 m. The numbers show that Mount Everest could fit inside Challenger Deep with more than 2 km to spare. Other deep trenches in the Pacific are Tonga, Kuril-Kamatcha, Philippine and Kermadec Trenches which all contain parts with depths greater than 10000 m.

The great depths of oceanic trenches result in a special environment like water pressures 1000 times higher than the ones in the surface, constant temperatures just above freezing and absence of light to sustain any photosynthesis. These conditions are believed to have developed unique habitats with an extraordinary abundance of a few highly specialized organisms.

Researchers who are interested in trench exploration and study have to face unique logistical and engineering challenges because of the extreme depth. This is why trench exploration to date is very limited and only three humans have visited the sea floor below 6000 m. Much of the information we have about trenches and their living organisms, are acquired from the two sampling campaigns in the 50s (the Danish Galathea and the Soviet Vitjaz) and from photographic expeditions and sea floor sampling that were carried out remotely from the deep. These attempts hinted the existence of previously unknown processes, species and ecosystems.

Although knowledge about trenches is limited and hard to acquire, scientists believe that trench exploration would play a significant role in our lives on land. By studying ocean trenches, scientists can get a better understanding about the physical process of subduction and the causes of natural disasters that are related to these, like earthquakes and tsunamis. Finally the study of the hadal organisms that have adapted their lives in this harsh habitat can hold the key to biological and biomedical advances. Researchers have already discovered microbes that live in deep-sea hydrothermal vents and consist potential new sources of anti-cancer drugs and antibiotics. [65], [46]

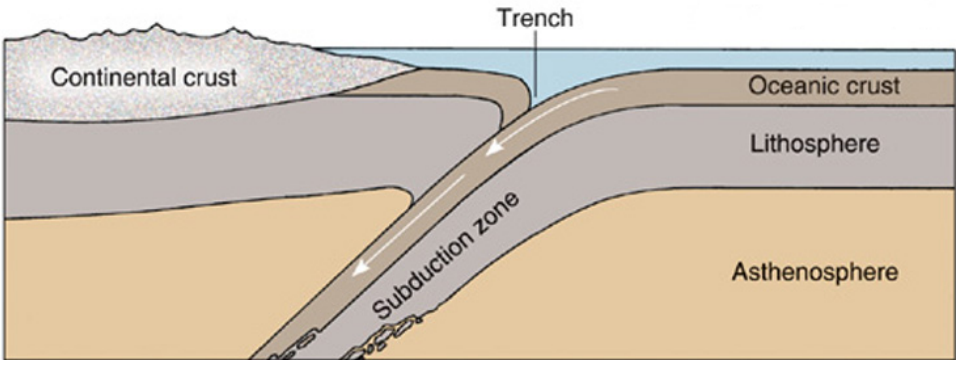


Figure 4.5: Subduction zone and trench formation [29]

4.3 Subduction Angle

When a plate subducts at a convergent boundary, it forms an angle with a horizontal plane called subduction angle or slab dip. Typically, the angles that are formed are steep but it is possible that some anomalies will occur like very shallow angles or extremely steep ones.

Seismologists use the hypocenter locations of both intermediate and deep earthquakes to represent the top surface of the descending slab. They came to the result that descending slabs form a characteristic dip along the entire length of a subduction arc, which is in an approximate depth range of 100 to 400 km.

There are four basic factors that affect the inclination of the subducted oceanic lithosphere and generally the subduction-zone geometry.

- Rapid upper plate motion towards the trench and active overriding of the subducting slab: The angle of subduction decreases for increased motion of the upper plate towards the trench.

- Rapid relative plate convergence: It works like the absolute motion of the upper plate, so the angle of subduction decreases for increased convergent rates. Luyendyk, in 1970, was the first one who suggested that there exists a common inverse relationship between the dip angle and the relative rate of convergence within the two plates.

- Anomalously low density of the oceanic lithosphere that subducts: This aspect includes the subduction of aseismic ridges, oceanic plateaus and intra-plate island-seamount chains. When the average density of the slab is reduced then the relative buoyancy of the lithosphere is reduced as well. This results in reduced subduction angles and it is common to find very low-angle subduction in these zones.

- Age of the subducting lithosphere: The younger the lithosphere, the more buoyant it is as it has lower density so it subducts at reduced angle.

The previous points reveal that a low-angle subduction is a result of the combination of rapid absolute upper-plate motion towards the trench, relatively rapid plate convergence, anomalously low-density of the subducting oceanic lithosphere and finally young subducting slab. On the other hand, normal or steeper subduction angles are a result of the combination of slower absolute upper mantle motion, slow relative plate convergence, old subducting lithosphere and of normal density.

Many scientists also suggest that the accretionary wedge, which is formed from sediment and slices of oceanic crust along the inner slope of the trench, can reduce the angle of subduction in the shallow part of the Benioff zone as it loads and depresses the subducting plate. Accretionary wedge and loading is though considered to be a subordinate factor that affects the dip angle compared to the previous four.

Some extreme examples of flat-angle subductions are in Central Chile at the Andean Volcanic Belt and in northern Peru where the dip angle is calculated to be 5 degrees. As for steep-angle subductions the Mariana Trench is notable with a dip angle around 81 degrees. [64], [10], [71]

4.4 Subduction Zones

There are many subduction zones all around the world as can be seen in figure 4.6. There are though some important zones that stand out among the others mostly because they produce earthquakes of high magnitude ($M \geq 7.5$) .

Table 4.1 contains information about the most distinguished and well known subduction regions across the earth. They are also shown on the map of figure 4.6. Age is measured in Ma which stands for Mega-annum (million years).

Region	Subd. rate (mm/yr)	Age of slab (Ma)	Date of max event	M
South Chile	70	26	1960	9.6
Central Chile	70	40	1922	8.3
New Zeland	43	120	1931	7.8
North Sumatra	33	72	2004	9.3
South Sumatra	51	61	1833	9.2
Nankai	57	20	1707	8.8
Kamchatka	78	84	1952	8.9
Kuriles	81	110	1963	8.5
Alaska	54	42	1964	9.3
West Aleutian	73	84	1965	8.7
Cascadia	42	9	1700	9.1
North Peru	63	29	1940	8.2
Mexico	30	4	1932	8.1
Tonga	185	120	1865	8.3

Table 4.1: Subduction zone convergence parameters and maximum earthquakes magnitude [56]

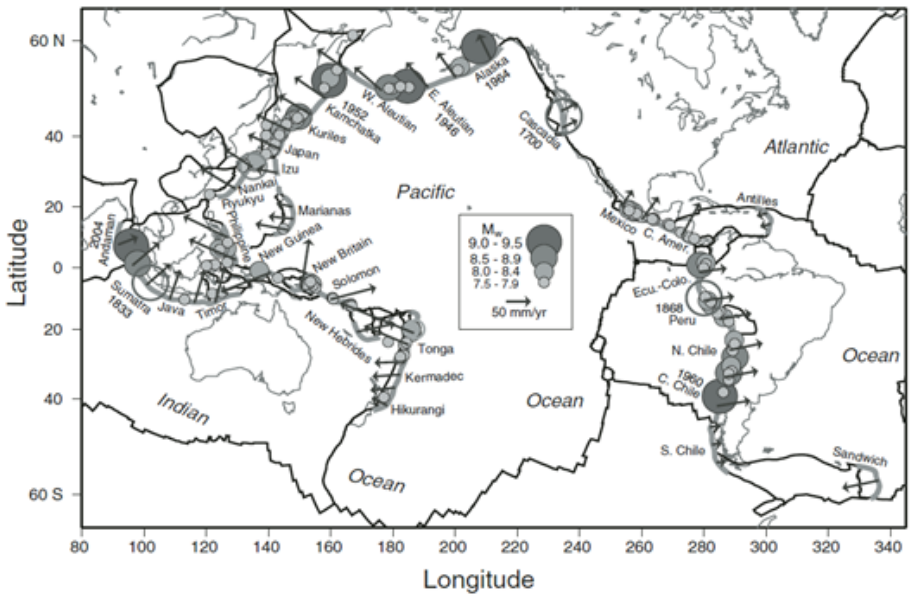


Figure 4.6: World's major subduction zones(thick grey lines) and tectonic plate boundaries. Filled circles show the locations of known earthquakes of $M \geq 7.5$ since 1900. Arrows show the horizontal velocity of subducting plate relative to overriding plate. [4]

Part II

The project

Chapter 5

Statistical Modelling of Subduction Zone Parameters

5.1 Subduction Zone Parameters

This chapter will examine the parameters that can be measured around the zones of subduction and will try to detect correlation within them which could indicate a relation. These relations will later be expressed as models with the help of simple or multiple linear regression. Studying subduction parameters and isolating possible relations within them, helps to better understand the origin and the extent of physical phenomena related to subductions, the development and behaviour of subduction zones and the relevant mantle dynamics. Equations within subduction parameters can firstly provide motivation for theoretical studies in the future and secondly can conceptually and kinematically join the geology of subduction zones and the plate tectonics.

In total 20 parameters will be examined, listed in Table 5.1 which are related to the geometry, the kinematics, the dynamics and the geology of 39 different subduction zones. For this purpose data will be used from Jaarard [23] which are shown in the Table 5.2. Following are some details about the parameters used, as given by the original author.

- **Arc Curvature**

  In nearly all cases, as an oceanic tectonic plate subducts, an arc-shaped island (when the overriding plate is oceanic crust) or an arc-shaped mountain belt (when the overriding plate is continental crust) is formed parallel to the oceanic trench. These arcs are often composed of volcanoes and are located on the overriding plate at such distance that the subducting slab is around 100 to 125 km deep. The curvature of the arcs, which is a result of Earth's sphericity, can be characterized by a radius of curvature when their geometry is treated like a part of a circle on the Earth's surface [12].

- **Slab - Benioff Zone Geometry**

  Benioff zones are located in subduction zones and are related to the

seismicity areas that are generated because of the downward motion of the subducting slabs. Locating and analysing the distributions of the earthquake hypocenters, gives an image of the geometry of the subducting slab. Table 5.2 contains information about the subducting slabs and their Benioff zones in respect to both length and dip amount .

Along the slab, the dip is not constant but it increases from the trench up to about 150 km depth. Below this depth it remains almost constant. For this reason, Table 5.2 contains three different measurements of the slab dip, each one for the three depths that normally big dip shifts occur. The shallow dip or DipS denotes the dip of the slab from the trench up to 60 km depth while the intermediate dip or DipI denotes the dip from the trench up to 100 km depth. Finally the deep dip or DipD contains the information of the slab dip within the depths of 150 and 400 km and DipU is the angle in which the slab descends into the mantle.

As for the length aspect, Table 5.2 contains data about the total length of the Benioff zone on the slab which is measured along its upper surface, the horizontal extent of the Benioff zone on the slab and finally there is the parameter of the maximum depth of the Benioff zone.

- **Convergence Rates and Absolute Motions**

  The way to determine the present convergent directions or else the obliquity of convergence at subduction zones, is through slip vectors from interplate thrust earthquakes. On the other hand, convergence rates cannot be measured directly but only through worldwide motion models based on calculations like focal mechanisms and transform fault azimuths. The convergence rates in Table 5.2 give the perpendicular to the trench convergent rates and are products of two different models by Chase in 1978 [6] and Minister and Jordan in 1978 [32]. In Table 5.2 the first model is denoted as C while the second as M. The convergent rate denoted as Vc describes the relative motion of the overriding plate and the subducting plate while the convergent rate denoted as Vcba describes the relative motion of the forearc and the subducting slab.

  The absolute motion parameters in Table 5.2 are also products of the two worldwide models of Chase and of Minister and Jordan as mentioned before. They are denoted by C or M and used the fixed-hot-spot hypothesis to calculate the absolute motions. The absolute motion describes the perpendicular to the trench motion of the overriding plate and when it has a positive sign, it means that the motion is towards the subducting plate.

- **Slab Age**

  In Table 5.2 there are two measures that describe the age of the slab. The first one is the average age of the slab that is now entering the

	Parameters	Symbols Used	Units
Slab	shallow dip (to 60 km depth)	DipS	deg
	intermediate dip (to 100 km depth)	DipI	deg
	deep dip (150-400 km)	DipD	deg
	descent angle of slab into mantle	DipU	deg
	horizontal extent of Benioff zone	Horizontal	km
	maximum depth of Benioff zone	SDepth	km
	length of Benioff zone on the slab	Length Slab	km
	slab age at trench	Age Slab	m.y.
	age of slab tip	Age Tip	m.y.
	time since slab tip subducted	T_x	m.y.
Trench	maximum trench depth	Tdepth	km
	relative trench depth	Dd	km
Upper Plate	duration of subduction	Age Arc	m.y.
	arc-trench gap	Gap (a-t)	km
	arc radius of curvature	RC	deg
Relative Motion	convergence rate	V_c	cm/yr
	convergence rate including back-arc spreading	V_{cba}	cm/yr
	absolute motion, overriding plate	V_{oa}	cm/yr
	rollback (absolute motion, forearc)	Rollback	cm/yr
	maximum cumulative earthquake moment	M	

Table 5.1: The parameters that will be examined with the symbols that will be used and their units

trench and the second one is the average age of the tip of the slab at the time that this slab tip entered the subduction zone. The age of the tip is calculated through the relationship:

$$A_t = A_s * L_s * (dA/dL - 1/V_{cba})$$

where A_s is the age of the crust that is now entering the trench, L_s is the downdip length of the slab, dA/dL (m.y/km) gives the age gradient of the subducting slab perpendicular to the trench (gives positive result if the age of the slab increases while it subducts) and finally V_{cba} (km/m.y.) is the convergent rate including the back-arc spreading. Finally Table 5.2 contains information about the time since the slab tip subducted, denoted as T_x.

- **Trench Depth**

  To define the trench depth, Table 5.2 uses two different measures, the maximum trench depth denoted by Tdepth and the relative trench depth denoted by Dd. Relative trench depth is the result of the difference between the maximum trench depth and the abyssal plain depth. Abyssal plains are the sediment-covered areas of the ocean floor which are formed on the top of a trench.

Name	Symbol	Slab											Trench	
		DipS	DipI	DipD	DipU		Horizontal	SDepth	Length Slab	Age slab	Age Tip	Tx	TDepth	Dd
					C	M								
Aegean	AEG	20	25	43	73	100	290	180	340			8,9		
Makran	MAK	8	12				430	80	440	97	87	11,9		
Andaman	AND	19	22				270	140	310	72	51	9,7		
Sumatra	SUM	16	19	50	50	48	370	170	380	55	45	6,1	5,9	0,7
Java	JAV	16	21	63	68	61	570	630	870	138	107	10,6	7,45	2,15
North Sulawesi	SUL	18	26	68			290	300	460	38	9	15,3		
Sangihe	SAN			56			620	670	920					
Philippine	PHL	43	41				100	130	170			3,7	10,06	4,1
Ryukyu	RYK	19	23	45	64	45	300	280	440	49	37	14,7	7,51	1,97
SW Japan	SJP	10					330	75	345	21	3	26,5	4,8	0,19
New Zealand	NZL	12	18	50	24	18	440	270	540	98	97	16,4		
Kermadec	KER	23	30	71	87	79	330	500	640	113	114	10,4	10,05	4,49
Tonga	TON	23	28	57	74	67	670	650	940	120	110	9,6	10,8	5,24
New Hebrides	NHB	36	44	73			190	280	330	52	45		7,07	2,56
Solomon	SOL	35	42	84	98	96	300	520	600	50	45	5	8,94	4,42
New Britain	NBR	30	35	58	76	70	250	290	390	50	47	2,7	8,24	3,73
Palau	PAL									31			8,05	3,55
Yap	YAP									33			8,53	
Marianas	MAR	19	24	81	87	66	310	700	860	155	134	9,4	9,66	4,95
Izu-Bonin	IZU	22	28	65	57	45	580	600	860	146	122	11,3	9,7	3,54
NE Japan	NJP	15	19	27	29	28	1340	600	1480	130	94	14,9	8	1,95
Kurile	KUR	22	28	50	56	49	650	600	890	119	89	10,2	9,78	4,27
Kamchatka	KAM	19	25	54	60	54	580	600	860	90	72	9,8	7,5	2
Central Aleutians	ALU	25	31	64	84	72	250	270	370	54	48	6,2	7,14	1,64
Alaska Peninsula	AKP	9	13	51	46	48	470	155	530	46	50	12,9		
Alaska Peninsula	ALA	7	10				620	160	650	46	48	10,3		
Cascades	CAS	9					97	34	103	8	9	3		
SW Mexico	WMX	19	25				210	90	230	15	14	3,9	5,29	1,79
SE Mexico	SMX	14	18	53	65	68	400	210	480	14	14	6,7	5,12	1,62
Middle America	NIC	30	38	65	59	63	170	210	280	23	23	4,3	6,66	2,56
Lesser Antilles	ANT	16	22	51	80	67	320	200	410	68	78	11,1	7	1,8
Colombia	COL	22	26	38	56	59	315	215	390	15	8	5,7	4	0,8
Ecuador	ECU		31	32				210		32				
Peru	PER	14	13	5			710	190	730	45	57	8,9	6,3	1,6
North Chile	NCH	20	21	30	35	43	810	600	1040	82	112	10,4	8,05	3,35
Central Chile	CCH	16	14	5			720	180	730	48	68	7,4	6,4	2,2
South Chile	SCH	13	16	30	36	42	490	170	520	26	30	5,4	4,7	0,6
Tierra del Fuego	TDF									20				
South Sandwich	SCO	31	38	67	116	111	220	250	350	49	74	4,6	8,26	3,81

Name	Symbol	Arc			Relative Motion								M
		Age Arc	Gap (a-t)	RC	Vc		Vcba		Voa		Rollback		
					C	M	C	M	C	M	C	M	
Aegean	AEG	11	210	2	0,7	0,4	3,8	3,4	-1,1	-0,1	2	2,9	
Makran	MAK	46	480		3,7	3,7	3,7	3,7	-0,3	-0,4	-0,3	-0,4	
Andaman	AND	100	270	7	2,1	1,4	3,2	2,6	-0,5	-0,1	0,6	1,1	7,9
Sumatra	SUM	27	300	22	6,2	5,6	6,2	5,6	0	-0,2	0	-0,2	7,1
Java	JAV	27	300	39	8,2	8,1	8,2	8,1	0,7	-0,3	0,7	-0,3	
North Sulawesi	SUL	7	150	6	3	3	3	3					
Sangihe	SAN	13		7					0,6	0			
Philippine	PHL	6	250	18	4,6	8	4,6	8	0,3	0,1	0,3	0,1	8
Ryukyu	RYK	55	150	14	3	4,8	3	4,8	1,1	0	1,1	0	8,6
SW Japan	SJP	175	300	4	1,3	2	1,3	2	1,2	0	1,2	0	7,8
New Zealand	NZL	21	280	50	3,3	3	3,3	3	-3,7	-5,3	-3,7	-5,3	8,1
Kermadec	KER	30	165	50	5,1	4,7	10,5	10,1	-2,5	-4	2,9	1,4	8,3
Tonga	TON	24	185	45	7,5	7,2	12,5	12,2	-1,2	-2,7	3,8	2,3	7,9
New Hebrides	NHB	8	140	30	8,8	8,4	9	8,6	6,3	6,7	6,8	6,9	
Solomon	SOL	8	100	12	12	11,8	12	11,8	3	2,5	3	2,5	
New Britain	NBR	8	125	6	4,3	4,5	14,7	14,9	-5,8	-7,1	4,6	3,3	
Palau	PAL	45		3	0,6	0,6	0,6	0,6	-3,2	-7,1	-3,2	-7,1	
Yap	YAP	45		4,5	1,2	1	1,2	1,2	-3,9	-8,1	-3,9	-8,1	7,2
Marianas	MAR	45	225	8	6	3,4	10,3	7,7	-3,3	-6,5	1	-2,2	7,2
Izu-Bonin	IZU	45	210	85	7,6	6,3	7,6	6,3	-1,2	-3	-1,2	-3	8,35
NE Japan	NJP	115	300	11	9,9	10,2	9,9	10,2	0,7	0,2	0,7	0,2	8,8
Kurile	KUR	82	190	15	8,7	8,6	8,7	8,6	1,1	-0,2	1,1	-0,2	9
Kamchatka	KAM	153	225	15	8,8	8,6	8,8	8,6	1	0	1	0	8,7
Central Aleutians	ALU	56	190	14	6	5,9	6	5,9	2,1	0,9	2,1	0,9	
Alaska Peninsula	AKP	160	400	14	4,1	3,9	4,1	3,9	-0,5	-0,3	-0,5	-0,3	9,1
Alaska Peninsula	ALA	160	470		6,3	6,1	6,3	6,1	1,2	0,8	1,2	0,8	
Cascades	CAS	175	280	90	3,4	3,4	3,4	3,4	2,2	1,8	2,2	1,8	
SW Mexico	WMX	90	240	15	5,9	5,6	5,9	5,6	1,9	2,1	1,9	2,1	
SE Mexico	SMX	90	380	15	7,2	7	7,2	7	1,6	1,9	1,6	1,9	8,4
Middle America	NIC	100	150	10	6,5	6,4	6,5	6,4	-0,8	-0,3	-0,8	-0,3	7,5
Lesser Antilles	ANT	48	260	5,5	3,7	3,7	3,7	3,7	1,8	1,1	1,8	1,1	8,8
Colombia	COL	242	270	173	6,8	6,9	6,8	6,9	2,5	2,9	2,5	2,9	
Ecuador	ECU	226	275	6	7,9	7,7	7,9	7,7	2,5	3	2,5	3	8,6
Peru	PER	226		19	8,2	7,7	8,2	7,7	1,5	2,7	1,5	2,7	8,65
North Chile	NCH	226	300	140	10	8,9	10	8,9	1,6	2,9	1,6	2,9	8,65
Central Chile	CCH	226		10	9,8	8,5	9,8	8,5	2	2,8	2	2,8	9,5
South Chile	SCH	226	300	55	9,7	8,2	9,7	8,2	1,7	2,5	1,7	2,5	
Tierra del Fuego	TDF	150	230	7	2,3	2,1	2,3	2,1	1,2	2,3	1,2	2,3	7
South Sandwich	SCO	30	150	3	0,9	0,9	7,9	7,9	-0,4	-1,1	6,6	5,9	

Table 5.2: All the 20 different subduction parameters for 39 subduction zones

5.2 Previous work

Dating back to the 80's, scientists are trying to isolate relationships among subduction parameters and apply multivariable statistical analysis to the existing datasets. At that point, they already had data concerning subduction parameters for many different subduction zones around the world. Many studies had already been conducted and each one of them could provide information about individual subduction parameters. Of course the use of statistical methods in order to determine correlations among subduction zone variables is used until today from scientists who use more recent or focused data for this purpose. The majority of the work done so far, covers targeted areas of interest like parameters that affect the seismicity in subduction zones, rather than relations among all the available parameters.

The first complete and most accurate attempt to gather all the existing subduction parameters' data for many subduction zones was conducted by Richard D. Jarrard with his paper "Relations Among Subduction Parameters" [23]. This study focuses on the determination of quantitative cause-effect equations, among different subduction parameters, applying stepwise regression analysis to some proportion of the dataset. Finally equations were formulated describing parameters like the trench relative depth, the Benioff zone length on the slab, the maximum earthquake moment , the arc-trench gap and the different slab dips. Most important among those, the age of the subducted slab combined with its relative convergence rate, was found to give great correlation with the length of the Benioff zone.

The majority of the other works that use subduction parameters to detect relations within them, are mostly focused on one attribute. An excellent example of this kind of work is the paper and study of Maria Sdrolias and R. Dietmar Muller [51]. They conducted a time and space dependent study focusing on the parameters that drive the mechanism of the back-arc basins development and extension at selected subduction zones. Another differentiation in this study is that the data used are grid data along the subduction zones instead of average values. They managed to prove that back-arc basins development depends on the age of the subducted lithosphere and on the angle of the subducting slab. Their analysis indicated that "the driving mechanism for back-arc extension is a combination of surface kinematics, properties of the down-going slab, the effect of lateral mantle flow on the slab, and mantle wedge dynamics". Hiroo Kanamori also focused on the back-arc basins and more specifically in the article "Back-arc opening and the mode of subduction", examines the factors that cause an active or not back-arc opening [68]. The author reached the conclusion that the two cases are caused "by either differences in the stage of subduction process (evolutionary model) or differences in the motion of the landward plate".

Another example of targeted focus is the paper of W.P. Schellart and N. Rawlinson titled "Global correlations between maximum magnitudes of subduction zone interface thrust earthquakes and physical parameters

of subduction zones" [49]. The focus of this paper is based on the observation that some subduction zones produce giant thrust earthquakes while others produce relatively small ones. For this reason they investigate if and which subduction zone parameters can affect the variability of seismicity around these areas. They also try to correlate maximum recorded moment magnitude (M) with these parameters. The results showed that big thrust earthquakes "occur for rapidly shortening to slowly extending overriding plates, slow trench velocities, moderate to high subduction partitioning ratios, low subduction thrust dip angles, low subduction thrust curvature and low trench curvature angles". Another study that focuses on the seismic zones around subduction zones is one by Heuret, Lallemand, Funiciello, Piromallo and Faccenna titled "Physical characteristics of subduction interface type seismogenic zones revisited" [20]. They created a database of data related to their subject based on global earthquake catalogues for the period between 1900 and 2007 and they tried to isolate correlations between parameters through statistics, to detect cause-effect relationships. The subduction velocity was found to be the most highly controlling parameter for large seismic rates. The relation between the structural features of fore-arc basins and large subduction zone earthquakes was studied by Llenos and McGuire in "Influence of fore-arc structure on the extent of great subduction zone earthquakes" [30]. They proved that the most great subduction zone ruptures occur where long-lived material heterogeneity occurs instead of short time-scale stress heterogeneities. It is the trench parallel gravity anomalies that determine the seismogenic behaviour along subduction zones. Peterson and Seno try to quantify the shallow seismic activity in subduction zones by calculating the seismic moment release rates and seismic slip rates [39]. In the article "Factors affecting seismic moment release rates in subduction zones", they evaluate 24 different subduction zones and they find that the factor which highly affects the strength of seismic coupling is the age of the subducting lithosphere and later the absolute velocity of the upper plate. They failed to find any relationship between the moment release rate and the convergence velocity. Driven by the thought that coupling within the down-going plate and the upper one is directly related to the earthquake size that a subduction zone produces Ruff and Kanamori used multivariate regression to find the physical subduction parameters that are correlated with coupling [47]. In their paper "Seismicity and the subduction process", they found that convergence rate and lithosphere age were the primary correlated parameters with the intraplate coupling.

Parameters that affect the slab dip and correlations within them were investigated by Serge Lallemand and Arnauld Heuret in their work titled "On the relationships between slab dip, back-arc stress, upper plate absolute motion, and crustal nature in subduction zones" [27]. For this purpose they applied statistical analysis of oceanic subduction zone parameters and they resulted in evidence that "the upper plate absolute motion plays an important role on slab dip, as well as on upper plate strain". Hager and O'connell studied also the relationship between flow in the mantle and subduction zone dip angles [18] . In their study "Subduction

zone dip angles and flow driven by plate motion" they created some flow models that can accurately predict the observed dips of subducted slabs when the flow extends into the lower part of the mantle. That means that the dip of a subducted slab is mostly determined by the large scale flow which is imposed by the movement of the plates. It is worth mentioning that in order to develop their kinematic models they ignored the buoyancy forces and took into account the observed plate geometries and velocities.

5.3 My work

This project is based on the data collection of Jarrard's paper, as was mentioned before in the chapter, and its purpose is to use statistical methods to reveal possible and hidden relations within many subduction parameters. Finding high correlations can indicate relations among these parameters, which is the objective of using mathematical tools. As an extent, these relations are modelled through mathematical equations.

In order to manipulate the data and extract information and relations within them, a programming language and software environment called R is used, which is suitable for statistical computing and graphics. R is used in a wide extent for statistical works, data mining and data analysis which coincides with this work's purposes. Finally, it provides numerous graphical front-ends as part of the command line interface. Except from R, the Excel spreadsheet application is shortly used for the calculation features and graphing tools that it provides.

The statistical methods that are later used and explained in detail are Pearson's product – moment correlation coefficient so that the possible correlations among our parameters are detected and they are depicted in panel plots. Then cluster analysis is applied to our data so that we group the similar parameters in clusters. Finally multiple regression is used to express the relationships between the subduction parameters.

5.3.1 Correlation

In statistics, correlation refers to the dependence or else the statistical relationship between two random variables and describes the degree of relationship between these variables.

There is a wide variety of types of correlations depending on the data that we measure. Pearson Product - Moment Correlation (PPMC) is best suitable when both variables are measured at an interval level. For two ordinal variables Spearman rank Order Correlation or Kendall rank order Correlation would be more appropriate. For other types of variables different type of correlations can prove to be much more suitable. The nature of our data indicates that Pearson Product - Moment Correlation is very suitable, and we will use Pearson correlation coefficient to investigate the relationships within our parameters.

Pearson correlation coefficient indicates the strength and the type of the linear relationship between two random variables and is denoted by the

symbol r. This coefficient takes values between -1.0 and +1.0. Negative values indicate a negative relationship between the variables and the closest the coefficient is to -1.0 the strongest the negative relationship. On the contrary if the coefficient has a positive value that indicates a positive relationship between the variables which becomes strongest when it is close to +1.0. Finally coefficient of value 0 indicates that there is no linear relationship between the two variables.

Pearson correlation coefficient can be calculated using the formula

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Figure 5.1 shows all the pairwise calculated correlation coefficients of our parameters and each image uses different velocity models. The dataset used for this study contains many missing values and for this reason pairwise deletion of the missing data is used. Mean substitution would change considerably the values of correlations as there are enough missing data and listwise deletion of missing data would be fruitless as all the parameters contain at least one missing value.

Figure 5.2 shows the pairwise scatter plots of all the parameters under evaluation, gathered in Panel Plots. Through these plots we can easily and optically detect some relationships within pairs of variables and the correlation data retrieved from the figure 5.1 should agree with the relationships that derive from the panel plots.

It is easy to observe that there are many pairs of variables that give high correlation results which is also obvious in their scatter plots. Some of these correlations though, especially the most outstanding ones with $r \geq 0.9$, are results of depending variables as some of them derive from others through the use of formulas. Examples of this case are the relationships between the parameters of Age Tip and Age Slab, Length Slab and Horizontal, DipI and DipS, Dd and TDepth. Excluding these correlations, the next highest ones are observed between the parameters of Gap (a-t) and DipI, Age Tip and TDepth, Age Slab and SDepth. It is also important to notice that the two different velocity models gave two big diversities in the correlation results. The first one is within the pair of parameters, DipD and DipU, where the Chase model gave a high correlation of $r = 0.74$ while the Minister and Jordan model gave a rather mediocre correlation of $r = 0.55$. The second diversity of the models is observed between the parameters' pair Rollback and Voa where the Chase model gave a low correlation of $r = 0.45$ while Minister and Jordan model gave the high correlation of $r = 0.71$

5.3.2 Clustering

After finding pairs of subduction parameters that give high correlation coefficients, one further step is performed. Groups of parameters that are related to each other in some way are determined and later quantified. In order to group the other variables that are close to each other, the technique of clustering is applied.

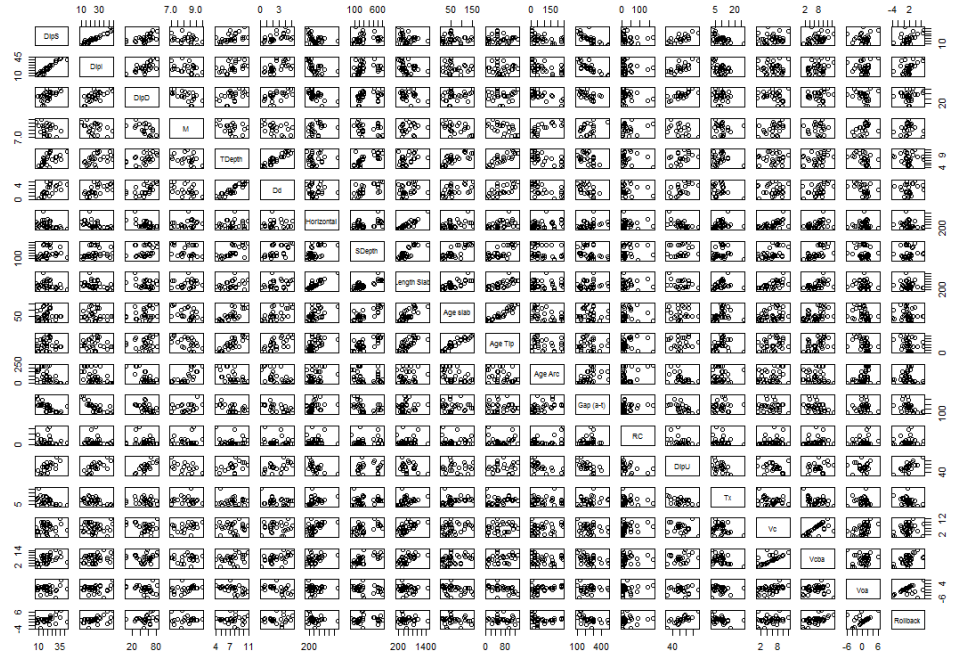
```
> cor(X, use="pairwise.complete.obs")
                  DipS        DipI        DipD           M      TDepth          Dd  Horizontal      SDepth Length Slab    Age slab
DipS          1.00000000  0.95913409  0.55592551 -0.29399595  0.47711724  0.56560555 -0.36894887  0.21441727 -0.18407007  0.03551833
DipI          0.95913409  1.00000000  0.66186269 -0.33280558  0.42248013  0.51813839 -0.52229632  0.17759764 -0.30871118 -0.06062703
DipD          0.55592551  0.66186269  1.00000000 -0.55908092  0.47105228  0.51959833 -0.58524478  0.34035650 -0.25103884  0.24921668
M            -0.29399595 -0.33280558 -0.55908092  1.00000000 -0.42209923 -0.39224443  0.26373701 -0.27403223  0.03973776 -0.48300193
TDepth        0.47711724  0.42248013  0.47105228 -0.42209923  1.00000000  0.91292141  0.11152435  0.68382523  0.39039218  0.73036972
Dd            0.56560555  0.51813839  0.51959833 -0.39224443  0.91292141  1.00000000 -0.03829704  0.61074470  0.26461841  0.55758025
Horizontal   -0.36894887 -0.52229632 -0.58524478  0.26373701  0.11152435 -0.03829704  1.00000000  0.50462093  0.90649526  0.48047665
SDepth        0.21441727  0.17759764  0.34035650 -0.27403223  0.68382523  0.61074470  0.50462093  1.00000000  0.80967467  0.80225238
Length Slab  -0.18407007 -0.30871118 -0.25103884  0.03973776  0.39039218  0.26461841  0.90649526  0.80967467  1.00000000  0.72385476
Age slab      0.03551833 -0.06062703  0.24921668 -0.48300193  0.73036972  0.55758025  0.48047665  0.80225238  0.72385476  1.00000000
Age Tip       0.05837705 -0.07529111  0.09776828 -0.50021322  0.81454003  0.67749291  0.47655847  0.73638356  0.69501563  0.92715848
Age Arc      -0.47159509 -0.46894930 -0.76036752  0.72637599 -0.59727677 -0.54046961  0.33379764 -0.26459053  0.12069042 -0.37489721
Gap (a-t)    -0.72805675 -0.81100536 -0.57281987  0.32517582 -0.49492179 -0.56271051  0.32149084 -0.33479871  0.08497347 -0.03895791
RC           -0.09125736 -0.10319183 -0.16495411  0.18220429 -0.14859396 -0.08172270  0.09843205  0.10949140  0.09223664  0.06113253
DipU          0.66221853  0.69281986  0.74044696 -0.47013689  0.36172912  0.49544373 -0.57707612  0.04913644 -0.34726376 -0.03374149
Tx           -0.48923594 -0.49939741 -0.08922587 -0.05125740 -0.03501354 -0.23643516  0.33217405  0.11325706  0.30228432  0.26371976
Vc            0.14537631  0.01272686 -0.22630706  0.41709213 -0.01833606  0.01646740  0.54400127  0.50465861  0.59967191  0.27901476
Vcba          0.38020765  0.27981508  0.05174828  0.08791446  0.28990809  0.44868088  0.35441168  0.61261689  0.52568097  0.39067396
Voa           0.04091197  0.07967658 -0.21296813  0.37534565 -0.45453043 -0.46915708  0.02803233 -0.18946843 -0.09091359 -0.31080187
Rollback      0.45340710  0.50064839  0.16911293 -0.04027418 -0.02848739  0.14936454 -0.20936217  0.02350233 -0.14282235 -0.10121705
                Age Tip     Age Arc   Gap (a-t)          RC        DipU          Tx          Vc        Vcba         Voa    Rollback
DipS         0.05837705 -0.47159509 -0.72805675 -0.09125736  0.66221853 -0.48923594  0.14537631  0.38020765  0.04091197  0.45340710
DipI        -0.07529111 -0.46894930 -0.81100536 -0.10319183  0.69281986 -0.49939741  0.01272686  0.27981508  0.07967658  0.50064839
DipD         0.09776828 -0.76036752 -0.57281987 -0.16495411  0.74044696 -0.08922587 -0.22630706  0.05174828 -0.21296813  0.16911293
M           -0.50021322  0.72637599  0.32517582  0.18220429 -0.47013689 -0.05125740  0.41709213  0.08791446  0.37534565 -0.04027418
TDepth       0.81454003 -0.59727677 -0.49492179 -0.14859396  0.36172912 -0.03501354 -0.01833606  0.28990809 -0.45453043 -0.02848739
Dd           0.67749291 -0.54046961 -0.56271051 -0.08172270  0.49544373 -0.23643516  0.01646740  0.44868088 -0.46915708  0.14936454
Horizontal   0.47655847  0.33379764  0.32149084  0.09843205 -0.57707612  0.33217405  0.54400127  0.35441168  0.02803233 -0.20936217
SDepth       0.73638356 -0.26459053 -0.33479871  0.10949140  0.04913644  0.11325706  0.50465861  0.61261689 -0.18946843  0.02350233
Length Slab  0.69501563  0.12069042  0.08497347  0.09223664 -0.34726376  0.30228432  0.59967191  0.52568097 -0.09091359 -0.14282235
Age slab     0.92715848 -0.37489721 -0.03895791  0.06113253 -0.03374149  0.26371976  0.27901476  0.39067396 -0.31080187 -0.10121705
Age Tip      1.00000000 -0.27773344 -0.01484682  0.07636893  0.03563964  0.13914288  0.23194770  0.41260114 -0.43014175 -0.15560335
Age Arc     -0.27773344  1.00000000  0.43858162  0.39287335 -0.53793787  0.03990632  0.30882296  0.02692208  0.39805951  0.01262915
Gap (a-t)   -0.01484682  0.43858162  1.00000000  0.20586654 -0.61085586  0.23021121 -0.03484499 -0.33284191  0.07356553 -0.44962437
RC           0.07636893  0.39287335  0.20586654  1.00000000 -0.37086308 -0.11868835  0.28857159  0.17204519  0.17736764  0.04552266
DipU         0.03563964 -0.53793787 -0.61085586 -0.37086308  1.00000000 -0.47412574 -0.30049914  0.22261674 -0.04620877  0.70931122
Tx           0.13914288  0.03990632  0.23021121 -0.11868835 -0.47412574  1.00000000 -0.24999136 -0.41546791 -0.06616168 -0.40467248
Vc           0.23194770  0.30882296 -0.03484499  0.28857159 -0.30049914 -0.24999136  1.00000000  0.73978621  0.48202935  0.25305047
Vcba         0.41260114  0.02692208 -0.33284191  0.17204519  0.22261674 -0.41546791  0.73978621  1.00000000  0.04422799  0.57018940
Voa         -0.43014175  0.39805951  0.07356553  0.17736764 -0.04620877 -0.06616168  0.48202935  0.04422799  1.00000000  0.45225069
Rollback    -0.15560335  0.01262915 -0.44962437  0.04552266  0.70931122 -0.40467248  0.25305047  0.57018940  0.45225069  1.00000000
>
```

(a) The velocities of the plates' movement calculated using the Chase model

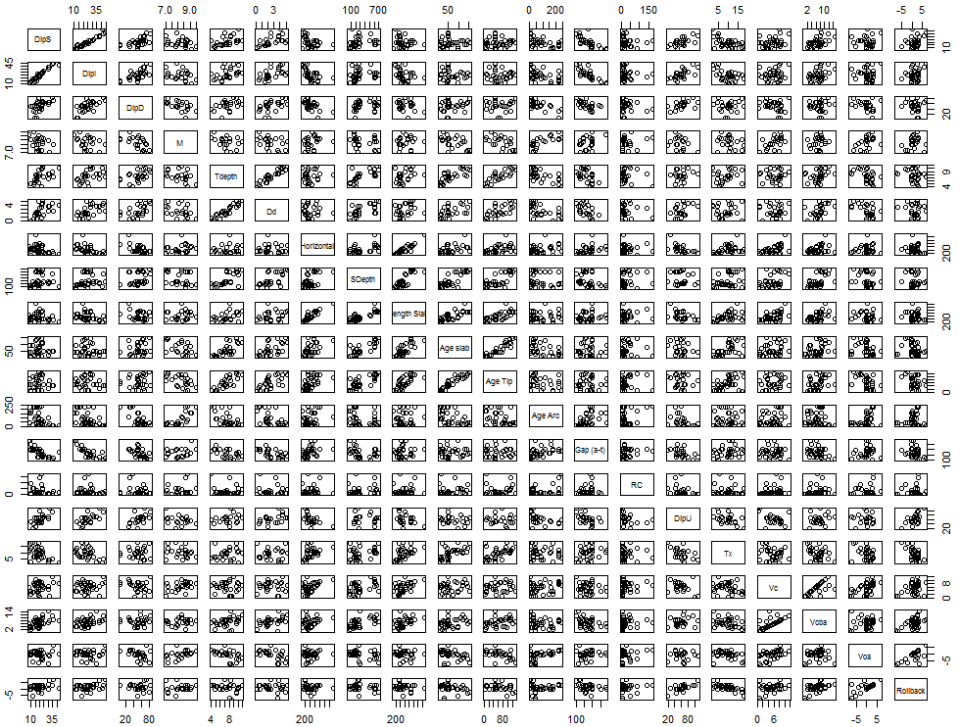
```
> cor(X, use="pairwise.complete.obs")
                  DipS        DipI        DipD           M       Tdepth          Dd  Horizontal      SDepth Length Slab    Age slab
DipS          1.00000000  0.959134087  0.55592551 -0.29399595  0.477117235  0.565605546 -0.36894887  0.2144173 -0.18407007  0.03551833
DipI          0.95913409  1.000000000  0.66186269 -0.33280558  0.422480128  0.518138385 -0.52229632  0.1775976 -0.30871118 -0.06062703
DipD          0.55592551  0.661862693  1.00000000 -0.55908092  0.471052283  0.519598334 -0.58524478  0.3403565 -0.25103884  0.24921668
M            -0.29399595 -0.332805581 -0.55908092  1.00000000 -0.422099233 -0.392244428  0.26373701 -0.2740322  0.03973776 -0.48300193
Tdepth        0.47711724  0.422480128  0.47105228 -0.42209923  1.000000000  0.912921413  0.11152435  0.6838252  0.39039218  0.73036972
Dd            0.56560555  0.518138385  0.51959833 -0.39224443  0.912921413  1.000000000 -0.03829704  0.6107447  0.26461841  0.55758025
Horizontal   -0.36894887 -0.522296322 -0.58524478  0.26373701  0.111524353 -0.038297039  1.00000000  0.5046209  0.90649526  0.48047665
SDepth        0.21441727  0.177597642  0.34035650 -0.27403223  0.683825226  0.610744702  0.50462093  1.0000000  0.80967467  0.80225238
Length Slab  -0.18407007 -0.308711180 -0.25103884  0.03973776  0.390392184  0.264618415  0.90649526  0.8096747  1.00000000  0.72385476
Age slab      0.03551833 -0.060627027  0.24921668 -0.48300193  0.730369719  0.557580246  0.48047665  0.8022524  0.72385476  1.00000000
Age Tip       0.05837705 -0.075291113  0.09776828 -0.50021322  0.814540026  0.677492908  0.47655847  0.7363836  0.69501563  0.92715848
Age Arc      -0.47159509 -0.468949301 -0.76036752  0.72637599 -0.597276775 -0.540469610  0.33379764 -0.2645905  0.12069042 -0.37489721
Gap (a-t)    -0.72805675 -0.811005359 -0.57281987  0.32517582 -0.494921790 -0.562710512  0.32149084 -0.3347987  0.08497347 -0.03895791
RC           -0.09125736 -0.103191834 -0.16495411  0.18220429 -0.148593963 -0.081722699  0.09843205  0.1094914  0.09223664  0.06113253
DipU          0.66685933  0.656191471  0.54800558 -0.32315138  0.201014257  0.421046584 -0.55713556 -0.1172183 -0.42618148 -0.22658437
Tx           -0.52219209 -0.511173547  0.01683186 -0.24772885  0.174958330  0.016990720  0.42415817  0.3286231  0.48788362  0.55718791
Vc            0.26641732  0.124401590 -0.23771346  0.48617457  0.007467368 -0.002094566  0.49067940  0.4276761  0.51476593  0.21468430
Vcba          0.51109257  0.401654378  0.05581749  0.12837096  0.341646431  0.475360929  0.30292773  0.5593701  0.45453384  0.34268848
Voa           0.01276414 -0.001455188 -0.38196785  0.45710950 -0.512715025 -0.513504286  0.04424009 -0.3238501 -0.14967173 -0.35419486
Rollback      0.36418879  0.348282016 -0.13083516  0.17614146 -0.240246716 -0.094380660 -0.14925163 -0.2040029 -0.21804984 -0.22486548
                Age Tip     Age Arc   Gap (a-t)          RC        DipU          Tx          Vc        Vcba         Voa    Rollback
DipS          0.05837705 -0.47159509 -0.72805675 -0.09125736  0.66685933 -0.52219209  0.266417316  0.51109257  0.012764137  0.36418879
DipI         -0.07529111 -0.46894930 -0.81100536 -0.10319183  0.65619147 -0.51117355  0.124401590  0.40165438 -0.001455188  0.34828202
DipD          0.09776828 -0.76036752 -0.57281987 -0.16495411  0.54800558  0.01683186 -0.237713457  0.05581749 -0.381967846 -0.13083516
M            -0.50021322  0.72637599  0.32517582  0.18220429 -0.32315138 -0.24772885  0.486174565  0.12837096  0.457109501  0.17614146
Tdepth        0.81454003 -0.59727677 -0.49492179 -0.14859396  0.20101426  0.17495833  0.007467368  0.34164643 -0.512715025 -0.24024672
Dd            0.67749291 -0.54046961 -0.56271051 -0.08172270  0.42104658  0.01699072 -0.002094566  0.47536093 -0.513504286 -0.09438066
Horizontal    0.47655847  0.33379764  0.32149084  0.09843205 -0.55713556  0.42415817  0.490679398  0.30292773  0.044240092 -0.14925163
SDepth        0.73638356 -0.26459053 -0.33479871  0.10949140 -0.11721826  0.32862313  0.427676112  0.55937005 -0.323850078 -0.20400290
Length Slab   0.69501563  0.12069042  0.08497347  0.09223664 -0.42618148  0.48788362  0.514765927  0.45453384 -0.149671730 -0.21804984
Age slab      0.92715848 -0.37489721 -0.03895791  0.06113253 -0.22658437  0.55718791  0.214684299  0.34268848 -0.354194865 -0.22486548
Age Tip       1.00000000 -0.27773344 -0.01484682  0.07636893 -0.12086408  0.46135186  0.128923640  0.33593488 -0.496417923 -0.33971215
Age Arc      -0.27773344  1.00000000  0.43858162  0.39287335 -0.40465005 -0.01377396  0.254466548 -0.03378197  0.476182898  0.22977014
Gap (a-t)    -0.01484682  0.43858162  1.00000000  0.20586654 -0.50290120  0.25926795 -0.055679200 -0.36912181  0.144292397 -0.28137600
RC            0.07636893  0.39287335  0.20586654  1.00000000 -0.30149007 -0.06371821  0.254704182  0.13855894  0.200523669  0.11676765
DipU         -0.12086408 -0.40465005 -0.50290120 -0.30149007  1.00000000 -0.51431213 -0.314955849  0.19958394  0.031360728  0.66427193
Tx            0.46135186 -0.01377396  0.25926795 -0.06371821 -0.51431213  1.00000000 -0.244138410 -0.35921903 -0.338077657 -0.60630635
Vc            0.12892364  0.25446655 -0.05567920  0.25470418 -0.31495585 -0.24413841  1.000000000  0.70991954  0.510193772  0.34001591
Vcba          0.33593488 -0.03378197 -0.36912181  0.13855894  0.19958394 -0.35921903  0.709919540  1.00000000  0.105366846  0.50658819
Voa          -0.49641792  0.47618290  0.14429240  0.20052367  0.03136073 -0.33807766  0.510193772  0.10536685  1.000000000  0.71428034
Rollback     -0.33971215  0.22977014 -0.28137600  0.11676765  0.66427193 -0.60630635  0.340015908  0.50658819  0.714280344  1.00000000
>
```

(b) The velocities of the plates' movement calculated using the Minister and Jordan model

Figure 5.1: Pairwise Pearson Product - Moment Correlation Coefficients for all the subduction parameters.



(a) The velocities of the plates' movement calculated using the Chase model



(b) The velocities of the plates' movement calculated using the Minister and Jordan model

Figure 5.2: Panel Plots for all the subduction parameters.

In general terms clustering addresses the problem of partitioning data points into groups based on their similarity and setting apart the dissimilar ones. In this study Agglomerative Hierarchical Clustering is used which is a bottom-up clustering technique where each cluster has sub-clusters and so on. Two measurements need to be defined before clustering data. Firstly the distance measure needs to be determined. This measure calculates the similarity or dissimilarity between two objects. There are various ways to define distance and some of them are the Pearson correlation or distances like euclidean, city block, supremum and Minkowski. Secondly a cluster method or linkage function needs to be defined. This method describes the way that two different objects will be clustered. The most common linkage functions are the single linkage, complete linkage, average linkage and ward linkage. A hierarchical clustering algorithm follows the next steps.

- Every single object (subduction parameter in this case) is considered to be a separate cluster.
- The clusters that minimize the linkage function when clustered together, are grouped and replaced by a combined cluster.
- The similarity between the new cluster and the rest of the objects is calculated using the same linkage method as before.
- This procedure stops when all the objects have been clustered.

The result is a binary dendrogram which shows all the steps of data grouping and whose root is the cluster that contains all the objects . The clustering tree represents the partitions' hierarchy throughout the whole process.

Figure 5.3 depicts the resulting cluster dendrograms of the subduction parameters using both Chase and Minister & Jordan velocity models, after agglomerative hierarchical clustering was applied. More specifically the distance measure used is the euclidean distance while the linkage method was complete linkage. The y-axis in the dendrogram, named height, is the value of the distance metric between the clusters.

In order to investigate the existence of clusters that do not appear because of the previous choices of euclidean measure and complete linkage, many other techniques were also tested. The combination of techniques that gave the most optically different results compared to the previous try, are the ones using euclidean distance and average linkage depicted in figure 5.4 and the ones using maximum distance and average linkage depicted in figure 5.5.

The optical result of the dendrograms in figures 5.3, 5.4, 5.5 is very different but this is not the case for the actual meaning of the dendrograms. If we examine the clusters formed, they may not look similar but the distances within them give the same grouping of data. The same happens with the two different velocity models but in a smaller extend. The clusters may look slightly different in the way that their children are ordered in some nodes, but in reality the results are identical.

If the trivial correlations are excluded, the main groups of parameters that will be later examined further are listed next.

- Group 1

  SDepth, Gap (a-t), Length Slab

- Group 2

  DipI, V_{oa} , Dd, Rollback, Tx, M, TDepth, Vc

- Group 3

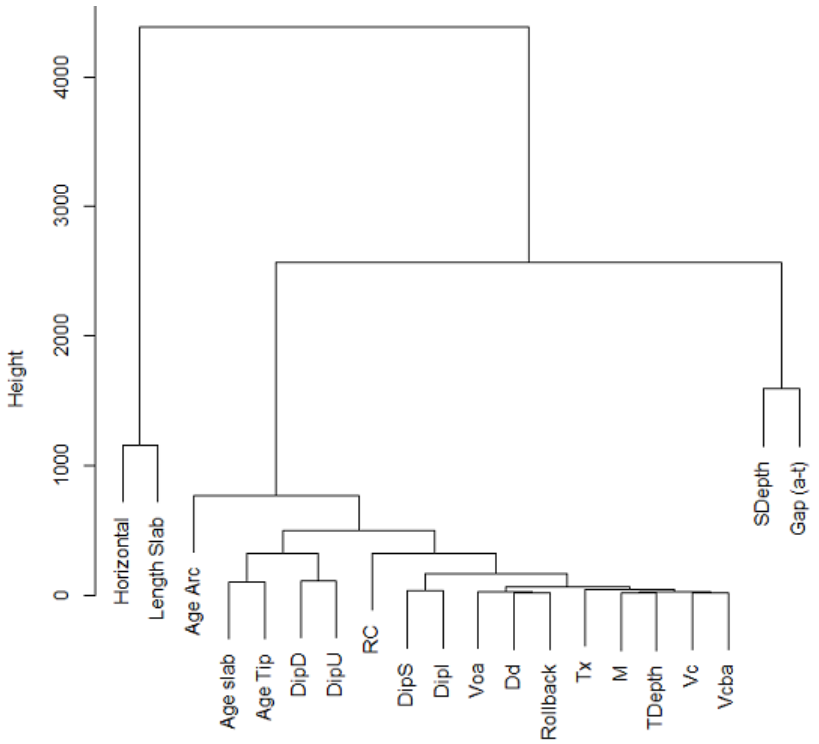
  Age Slab, DipD, DipU

5.3.3 Multiple Linear Regression

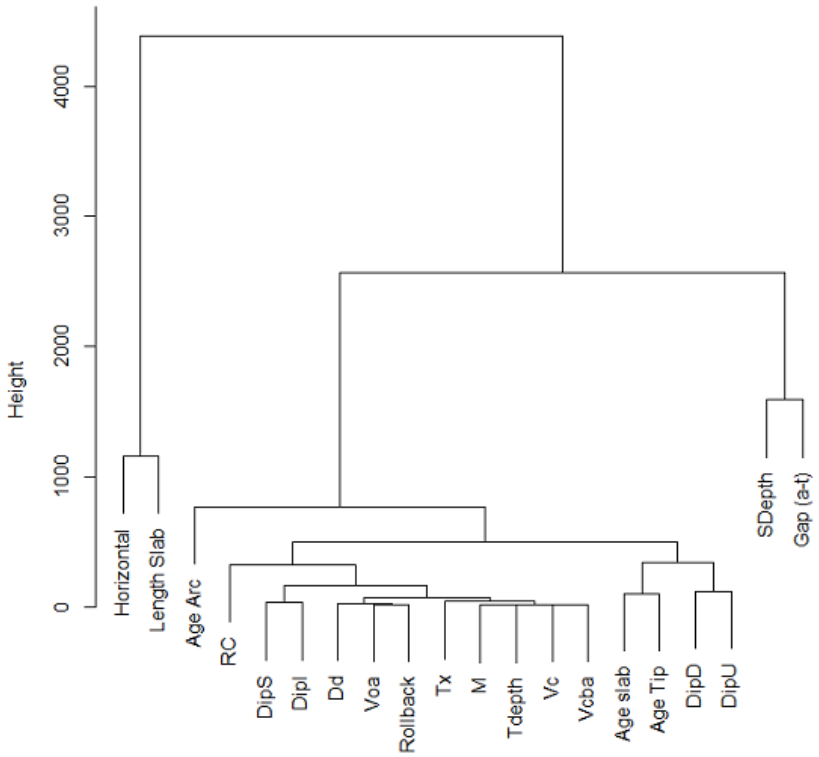
Linear regression is a statistical approach used to model the relationship between variables by fitting a linear equation to the observed data. The case of the relationship between a dependent variable and a single explanatory one is called simple linear regression while for more explanatory variables the process is called multiple linear regression. The previously presented scatter plots, correlation coefficients and clustering dendrograms have helped to determine the strength of the relationships within the subduction variables. In the next paragraphs multiple linear regression is used to model the relationships within the subduction variables.

A special case of multiple linear regression called stepwise regression, is also used. Stepwise regression is a semi-automated process which builds a model by either successively adding or removing (or both) variables in order to improve the t-statistics of their estimated coefficients in each step. This technique is useful when the independent variables are many and gives more power and information compared to the ordinary multiple regression.

Variables that are not instantly and physically measured by the geoscientists, like the arc radius of curvature or the convergence rates and the motions of the plates are used as explanatory variables for the models. The dependent variables are the length of the slab, the maximum cumulative earthquake moment, the intermediate and deep dip of the slab, the gap between arc and trench, the maximum depth of the Benioff zone on the slab and finally the relative trench depth. There is one main reason for modelling these specific dependent variables. Usually it is either impossible for scientists to measure them or it requires great resources of technology, money and manpower. Instead of that, they can be predicted and analysed using knowledge that is easily retrieved. Modelling the maximum earthquake moment for example, can give insight to why and when big earthquakes occur and in general we can get a better understanding of the subduction mechanism.

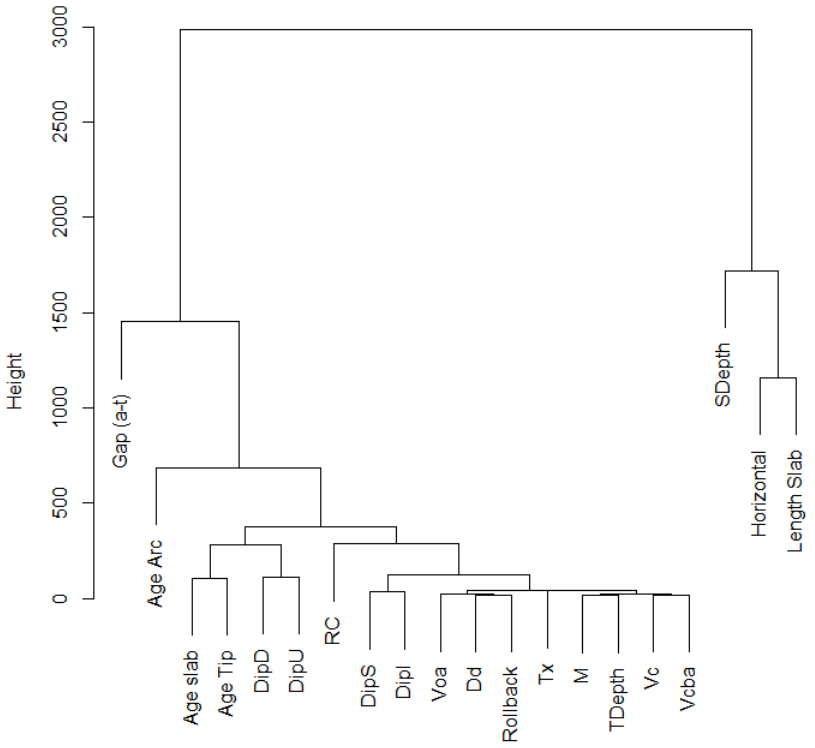


(a) The velocities of the plates' movement calculated using the Chase model

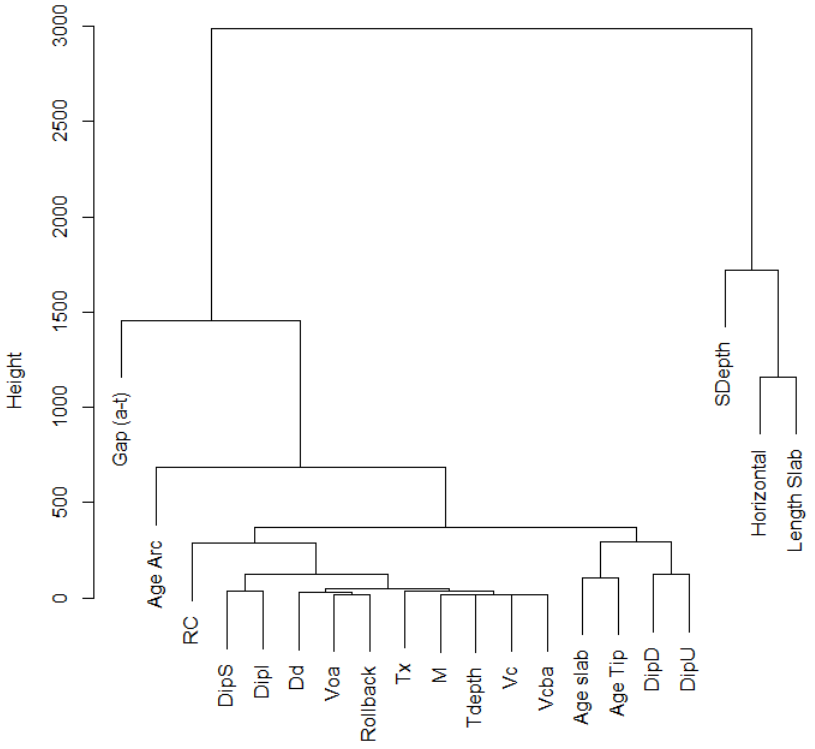


(b) The velocities of the plates' movement calculated using the Minister and Jordan model

Figure 5.3: Cluster Dendrograms of all the subduction parameters using euclidean distance and complete linkage.

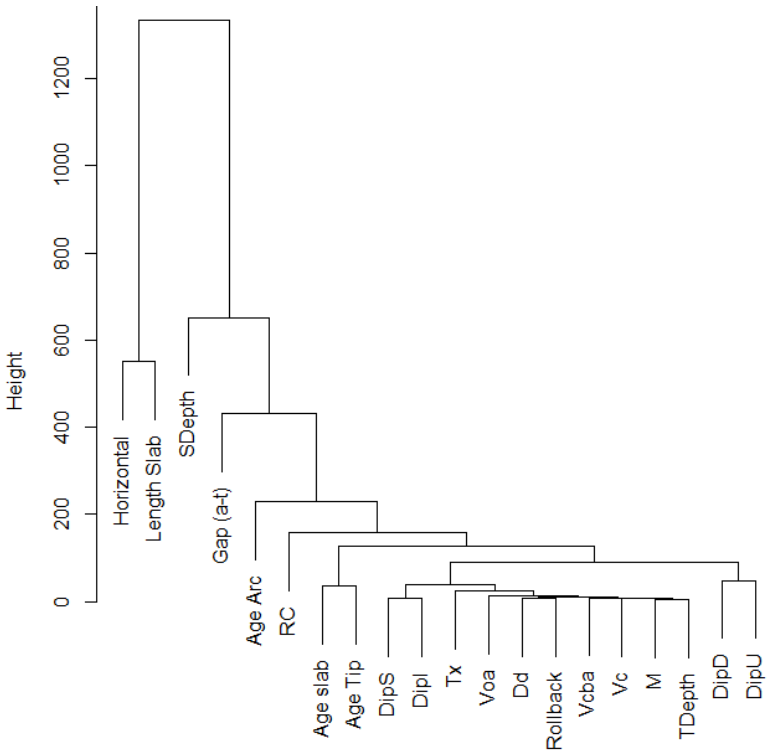


(a) The velocities of the plates' movement calculated using the Chase model

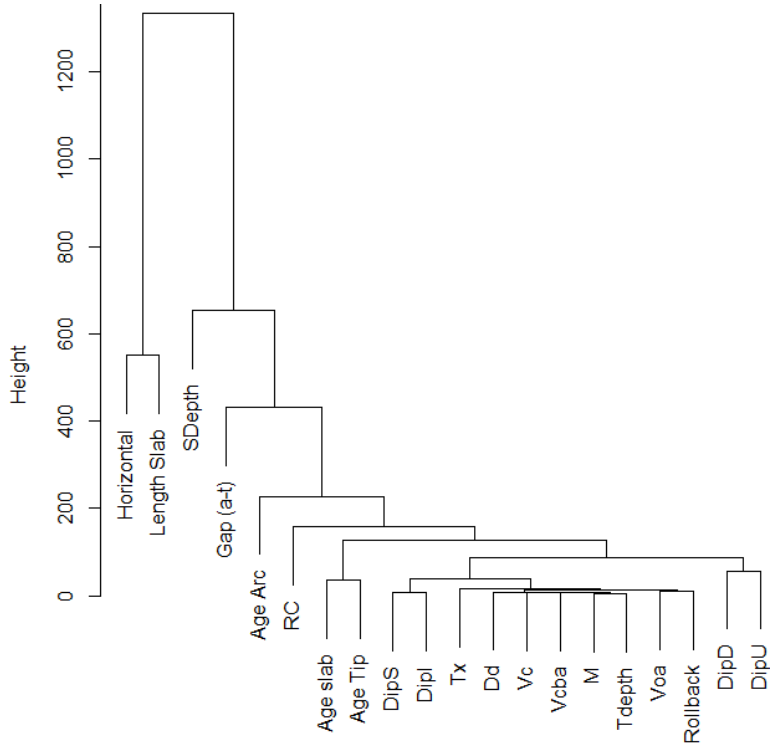


(b) The velocities of the plates' movement calculated using the Minister and Jordan model

Figure 5.4: Cluster Dendrograms of all the subduction parameters using euclidean distance and average linkage.



(a) The velocities of the plates' movement calculated using the Chase model



(b) The velocities of the plates' movement calculated using the Minister and Jordan model

Figure 5.5: Cluster Dendrograms of all the subduction parameters using maximum distance and average linkage.

- Intermediate Dip - DipI Model

  In order to create a model for the intermediate dip of the subducting slab by using multiple linear regression, we need to look at the correlation coefficients, the scatter plots and the clusters that are related to this parameter. The parameters that seem to have strong relationship with this variable and indicate that affect it strongly are the DipD, Gap (a-t) and DipU. These three parameters were the predictors into the first model as it seems in figure 5.6a. On the first step DipD had to be rejected as least significant, since it gave high p-value which is translated in high probability that this variable is not relevant with DipI. The second model uses two explanatory variables which both give low p-values and is shown in figure 5.6b. To be more precise, gap (a-t) variable gives extremely low p-value while DipU gives a value close to alpha level which is set at 0.05 (a commonly used level that gives 95 % confidence that the analysis is correct). The solution for deciding if DipU is a significant parameter to this model is to remove it and check the results. Figure 5.6c shows the last model which is under evaluation. Gap (a-t) variable seems to have a really low p-value here which means that is a significant one for our model but the overall results of the model seem to be much worse than before. R squared values which is an evaluation metric for the goodness of fit of the model decreased sharply. Both multiple R-squared value decreased from 0.8193 to 0.6577 and adjusted R-squared value decreased from 0.8012 to 0.6459. These results indicate that the best model for the dependent variable DipI includes both Gap (a-t) and DipU as explanatory ones. In order to verify this results backwards stepwise regression was used as depicted in figure 5.6d. The result of this process confirms that both Gap (a-t) and DipU should be used as explanatory variables and finally the model formed is:

  $$DipI = 36.64 - 0.07Gap(a-t) + 0.082DipU$$

  The previous model uses as one of the predictors, the DipU of Chase's velocity model, so it should also be analysed and expressed using the Minister & Jordan's velocity model. Figure 5.7 shows all the steps to the final model. The difference now is that the predictor DipU has p-value lower than the alpha level so there is no need to check a third model by excluding DipU from it. The differences between this model and the previous one is that the R-squared values are now increased to about 82 % compared to the previous 80 % and the coefficients are slightly changed. The new model becomes:

  $$DipI = 36 - 0.07Gap(a-t) + 0.097DipU$$

- Deep Dip - DipD Model

  Looking at the correlation coefficients, as well as the scatter plots and the clusters, between the DipD variable and the rest of the subduction parameters, four parameters distinguish as explanatory ones for the

```
Call:
lm(formula = DipI ~ DipD + Gap + DipU, data = data.subset1)

Residuals:
   Min     1Q Median     3Q    Max
-7.631 -1.869  0.025  2.106  6.270

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 34.58334    5.25568   6.580 2.68e-06 ***
DipD         0.06777    0.07196   0.942    0.358
Gap         -0.06856    0.01155  -5.934 1.03e-05 ***
DipU         0.05127    0.05170   0.992    0.334
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.37 on 19 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.8274,    Adjusted R-squared:  0.8001
F-statistic: 30.35 on 3 and 19 DF,  p-value: 1.879e-07
```

(a) First assessed model of M.L.R. with three explanatory parameters

```
Call:
lm(formula = DipI ~ Gap + DipU, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-8.3643 -1.9620  0.0887  1.8980  7.0441

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 36.63617    4.76893   7.682 2.17e-07 ***
Gap         -0.06999    0.01142  -6.128 5.47e-06 ***
DipU         0.08167    0.04028   2.027   0.0562 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.36 on 20 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.8193,    Adjusted R-squared:  0.8012
F-statistic: 45.34 on 2 and 20 DF,  p-value: 3.712e-08
```

(b) Second assessed model of M.L.R. with two explanatory parameters

```
Call:
lm(formula = DipI ~ Gap, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-10.005  -2.759  -1.258   2.114  15.492

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 44.24922    2.64552  16.726  < 2e-16 ***
Gap         -0.07496    0.01004  -7.465 3.16e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.205 on 29 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared:  0.6577,    Adjusted R-squared:  0.6459
F-statistic: 55.73 on 1 and 29 DF,  p-value: 3.157e-08
```

(c) Final assessed model of M.L.R.

```
Start:  AIC=59.49
DipI ~ DipD + Gap + DipU

       Df Sum of Sq    RSS    AIC
- DipD  1     10.07 225.81 58.537
- DipU  1     11.17 226.91 58.648
<none>              215.74 59.487
- Gap   1    399.80 615.54 81.601

Step:  AIC=58.54
DipI ~ Gap + DipU

       Df Sum of Sq    RSS    AIC
<none>              225.81 58.537
- DipU  1     46.41 272.22 60.836
- Gap   1    424.01 649.82 80.848

Call:
lm(formula = DipI ~ Gap + DipU, data = data.subset1)

Coefficients:
(Intercept)          Gap         DipU
   36.63617     -0.06999      0.08167
```

(d) Applying backwards stepwise regression to confirm the previous result

Figure 5.6: Steps of modelling the Intermediate Slab Dip using Multiple Linear Regression with Chase's velocity model

```
Call:
lm(formula = DipI ~ DipD + Gap + DipU, data = data.subset2)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2212 -1.7341 -0.2476  1.8691  6.0125

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 32.60462    4.87577   6.687 2.15e-06 ***
DipD         0.06642    0.05631   1.180   0.2528
Gap         -0.06645    0.01031  -6.447 3.52e-06 ***
DipU         0.07988    0.03794   2.106   0.0488 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.112 on 19 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.8528,    Adjusted R-squared:  0.8295
F-statistic: 36.68 on 3 and 19 DF,  p-value: 4.194e-08
```

(a) First assessed model of M.L.R. with three explanatory parameters

```
Call:
lm(formula = DipI ~ Gap + DipU, data = data.subset2)

Residuals:
    Min      1Q  Median      3Q     Max
-6.9046 -1.8616  0.4526  1.7794  6.3456

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 36.119577   3.896676   9.269 1.12e-08 ***
Gap         -0.070596   0.009783  -7.216 5.52e-07 ***
DipU         0.097209   0.035323   2.752   0.0123 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.142 on 20 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.842,     Adjusted R-squared:  0.8262
F-statistic: 53.29 on 2 and 20 DF,  p-value: 9.699e-09
```

(b) Second assessed model of M.L.R. with two explanatory parameters

```
Start:  AIC=55.82
DipI ~ DipD + Gap + DipU

       Df Sum of Sq    RSS    AIC
- DipD  1     13.47 197.45 55.450
<none>              183.98 55.825
- DipU  1     42.93 226.91 58.648
- Gap   1    402.43 586.41 80.486

Step:  AIC=55.45
DipI ~ Gap + DipU

       Df Sum of Sq    RSS    AIC
<none>              197.45 55.450
- DipU  1     74.77 272.22 60.836
- Gap   1    514.12 711.57 82.935

Call:
lm(formula = DipI ~ Gap + DipU, data = data.subset2)

Coefficients:
(Intercept)          Gap         DipU
   36.11958     -0.07060      0.09721
```

(c) Applying backwards stepwise regression to confirm the previous result

Figure 5.7: Steps of modelling the Deep Slab Dip using Multiple Linear Regression with Minister & Jordan's velocity model

```
Call:
lm(formula = DipD ~ DipI + Gap + Horizontal + AgeArc, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-18.761  -6.159  -1.230   7.848  23.116

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 31.438794  23.158133   1.358  0.18972
DipI         1.017111   0.477898   2.128  0.04593 *
Gap          0.050224   0.050845   0.988  0.33506
Horizontal  -0.015816   0.008983  -1.761  0.09360 .
AgeArc      -0.096949   0.032273  -3.004  0.00701 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.634 on 20 degrees of freedom
  (14 observations deleted due to missingness)
Multiple R-squared:  0.6588,    Adjusted R-squared:  0.5906
F-statistic: 9.654 on 4 and 20 DF,  p-value: 0.0001623
```

(a) First assessed model of M.L.R. with four explanatory parameters

```
Call:
lm(formula = DipD ~ DipI + Horizontal + AgeArc, data = data.subset1)

Residuals:
     Min       1Q   Median       3Q      Max
-19.9731  -5.0441   0.4371   6.8152  23.1378

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 49.68121   11.95699   4.155 0.000383 ***
DipI         0.77563    0.32428   2.392 0.025328 *
Horizontal  -0.01656    0.01024  -1.618 0.119366
AgeArc      -0.11776    0.03227  -3.650 0.001337 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.99 on 23 degrees of freedom
  (12 observations deleted due to missingness)
Multiple R-squared:  0.7268,    Adjusted R-squared:  0.6911
F-statistic: 20.39 on 3 and 23 DF,  p-value: 1.131e-06
```

(b) Second assessed model of M.L.R. with three explanatory parameters

```
Call:
lm(formula = DipD ~ DipI + AgeArc, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-18.365  -5.627  -0.425   6.827  25.201

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 39.98560    9.17718   4.357 0.000197 ***
DipI         0.91536    0.28966   3.160 0.004096 **
AgeArc      -0.13679    0.02897  -4.721 7.67e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.19 on 25 degrees of freedom
  (11 observations deleted due to missingness)
Multiple R-squared:  0.7029,    Adjusted R-squared:  0.6792
F-statistic: 29.58 on 2 and 25 DF,  p-value: 2.573e-07
```

(c) Final assessed model of M.L.R. with two explanatory parameters

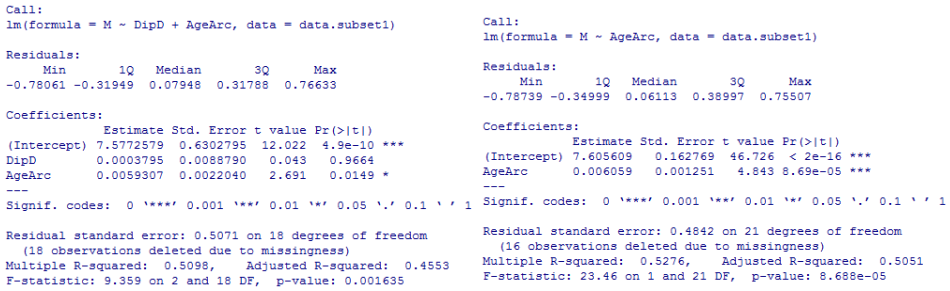
Figure 5.8: Steps of modelling the Deep Slab Dip using Multiple Linear Regression

model of DipD. These four parameters are DipI, Gap (a-t), Horizontal and AgeArc. As shown in figure 5.8a they are all part of the first model the assessment of which indicates that the variable gap (a-t) should be excluded as it gives a really high p-value meaning that it is not significant for the model. Removing this variable, a second model with three explanatory variables is generated and assessed as shown in figure 5.8b. In this step it is derived that the explanatory parameter Horizontal is not significant as it gives a p-value quite higher than the alpha value. That means that the next model to be assessed will have one less explanatory variable as shown in figure 5.8c. In this model both variables have low p-values and the overall model has high R-squared values at about 70 %. The final model that derives from all the above steps is:

$$DipD = 40 + 0.9DipI - 0.14AgeArc$$

- Maximum Cumulative Earthquake Moment - M Model

  As before, the first step demands to check the correlation coefficients, the scatter plots and the clusters within the dependent variable M and the rest of the subduction parameters. This results in two variables, DipD and AgeArc, that seem to affect M variable the most. As shown in figure 5.9a both variables were used for the first model to be assessed. This resulted in a high p-value for the DipD variable which



(a) First assessed model of M.L.R. with two explanatory parameters

(b) Final assessed model of M.L.R. with one explanatory parameter

Figure 5.9: Steps of modelling the Maximum Cumulative Earthquake Moment using Multiple Linear Regression

indicates that a new model should be created where DipD would be excluded. In figure 5.9b the second model is assessed by using only one predictor. AgeArc has a really low value indicating that it is relevant with M. As for the R squared values, they are quite high but not as high as the ones of the previous models. In this case the model seems to explain about 50 % of the variability in the response variable while in the previous models it explained about 80 % and 70 %. The final model that best explains M variable is:

$$M = 7.6 + 0.01 AgeArc$$

- Relative Trench Depth - Dd Model

  Looking at the correlation coefficients, the clusters and the scatter plots, DipI and AgeSlab can easily be recognised as the predictors that highly affect the dependent variable Dd . Figure 5.10a shows the first model including DipI and AgeSlab. It seems that they are both significant for this model as they both give low p-values. That quickly indicates a model with the R squared value being comparable to the one of the M model. In order to verify this model, backwards stepwise regression was used, as shown in figure 5.10b resulting in the same model which is:

$$Dd = -0.83 + 0.08 DipI + 0.02 AgeSlab$$

- Gap Between Arc & Trench - Gap (a-t) Model

  The predictors to be used for the first model of the dependent variable Gap (a-t) as resulted from the correlation coefficients, the clusters and the scatter plots are DipI and Dd. In figure 5.11a the first model assessment is shown and it is easy to notice that variable Dd has a high p-value which means that it is an insignificant for the model variable. That leads to a second model where Dd is excluded as shown in figure 5.11b. The variable DipI is highly significant in this

```
Call:
lm(formula = Dd ~ DipI + AgeSlab, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-1.2986 -0.8375  0.1210  0.7273  1.5352

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.834876   0.704274  -1.185  0.24908
DipI         0.083670   0.022507   3.718  0.00127 **
AgeSlab      0.018307   0.004299   4.258  0.00035 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9147 on 21 degrees of freedom
  (15 observations deleted due to missingness)
Multiple R-squared:   0.59,     Adjusted R-squared:  0.5509
F-statistic: 15.11 on 2 and 21 DF,  p-value: 8.602e-05
```

(a) First assessed model of M.L.R. with two explanatory parameters

```
Start:  AIC=-1.48
Dd ~ DipI + AgeSlab

          Df Sum of Sq    RSS     AIC
<none>                 17.570 -1.4843
- DipI     1    11.563 29.134  8.6521
- AgeSlab  1    15.172 32.742 11.4546

Call:
lm(formula = Dd ~ DipI + AgeSlab, data = data.subset1)

Coefficients:
(Intercept)          DipI      AgeSlab
   -0.83488      0.08367      0.01831
```

(b) Applying backwards stepwise regression to confirm the previous result

Figure 5.10: Steps of modelling the Relative Trench Depth using Multiple Linear Regression

```
Call:
lm(formula = Gap ~ DipI + Dd, data = data.subset1)

Residuals:
     Min       1Q   Median       3Q      Max
-109.100  -19.307   -6.703   13.894  118.471

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  414.612     34.944  11.865 1.66e-10 ***
DipI          -5.907      1.381  -4.277 0.000368 ***
Dd            -9.973      8.096  -1.232 0.232300
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 46.16 on 20 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.625,     Adjusted R-squared:  0.5875
F-statistic: 16.67 on 2 and 20 DF,  p-value: 5.5e-05
```

(a) First assessed model of M.L.R. with two explanatory parameters

```
Call:
lm(formula = Gap ~ DipI, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-120.792  -35.792   -9.566   18.595  137.138

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  472.592     31.937  14.798 4.76e-15 ***
DipI          -8.774      1.175  -7.465 3.16e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56.31 on 29 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared:  0.6577,    Adjusted R-squared:  0.6459
F-statistic: 55.73 on 1 and 29 DF,  p-value: 3.157e-08
```

(b) Final assessed model of M.L.R. with one explanatory parameter

Figure 5.11: Steps of modelling the Gap Between Arc & Trench using Multiple Linear Regression

model as it has high p-value while the model's R-value is about 65 %. The model of Gap (a-t) dependent variable is:

$$Gap(a-t) = 472.6 - 8.8DipI$$

- Length of Benioff Zone on the Slab - Length Slab Model

  As in all the previous cases, the first step is to check the scatter plots, the clusters and the correlation coefficient table to find the variables that best fit for predictors in this model. Figure 5.12a shows the first model including as explanatory parameters SDepth, AgeSlab and V_c. The results show that SDepth is a non significant variable as it takes high p-value. In figure 5.12b a new model is created by removing SDepth and keeping the other two predictors. Both of them seem to be significant values as they take really low p-values and the R-squared value of the overall model is high explaining about 70 % of the variability in the model of the length of the slab. The final model is :

$$LengthSlab = 4 + 4.3AgeSlab + 46V_c$$

```
Call:
lm(formula = LengthSlab ~ SDepth + AgeSlab + Vc, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-286.66 -101.77  -25.70   72.68  472.80

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  29.2186    76.7102   0.381   0.7062
SDepth        0.3436     0.2773   1.239   0.2256
AgeSlab       3.0656     1.1879   2.581   0.0154 *
Vc           37.7090    12.3109   3.063   0.0048 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 154.9 on 28 degrees of freedom
  (7 observations deleted due to missingness)
Multiple R-squared:  0.7376,    Adjusted R-squared:  0.7095
F-statistic: 26.24 on 3 and 28 DF,  p-value: 2.757e-08
```

```
Call:
lm(formula = LengthSlab ~ AgeSlab + Vc, data = data.subset1)

Residuals:
    Min      1Q  Median      3Q     Max
-305.37 -107.02  -32.09   88.89  459.98

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.0604    74.6535   0.054    0.957
AgeSlab       4.2761     0.6819   6.271 7.59e-07 ***
Vc           46.4717    10.1693   4.570 8.37e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 156.3 on 29 degrees of freedom
  (7 observations deleted due to missingness)
Multiple R-squared:  0.7233,    Adjusted R-squared:  0.7042
F-statistic: 37.89 on 2 and 29 DF,  p-value: 8.132e-09
```

(a) First assessed model of M.L.R. with three explanatory parameters

(b) Final assessed model of M.L.R. with two explanatory parameters

Figure 5.12: Steps of modelling the Slab Length using Multiple Linear Regression with Chase's velocity model

```
Call:
lm(formula = LengthSlab ~ SDepth + AgeSlab + Vc, data = data.subset2)

Residuals:
    Min      1Q  Median      3Q     Max
-294.02  -89.67  -21.42   82.56  416.68

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.9286    78.9241  -0.012  0.99070
SDepth        0.2720     0.2773   0.981  0.33506
AgeSlab       3.5102     1.2023   2.920  0.00685 **
Vc           43.6028    12.9991   3.354  0.00230 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 151.2 on 28 degrees of freedom
  (7 observations deleted due to missingness)
Multiple R-squared:  0.7501,    Adjusted R-squared:  0.7234
F-statistic: 28.02 on 3 and 28 DF,  p-value: 1.404e-08
```

```
Call:
lm(formula = LengthSlab ~ AgeSlab + Vc, data = data.subset2)

Residuals:
    Min      1Q  Median      3Q     Max
-309.48  -96.09  -17.85  104.69  397.07

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -25.320     74.855  -0.338    0.738
AgeSlab        4.502      0.650   6.927 1.30e-07 ***
Vc            51.273     10.377   4.941 2.99e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 151.1 on 29 degrees of freedom
  (7 observations deleted due to missingness)
Multiple R-squared:  0.7415,    Adjusted R-squared:  0.7237
F-statistic:  41.6 on 2 and 29 DF,  p-value: 3.017e-09
```

(a) First assessed model of M.L.R. with three explanatory parameters

(b) Final assessed model of M.L.R. with two explanatory parameters

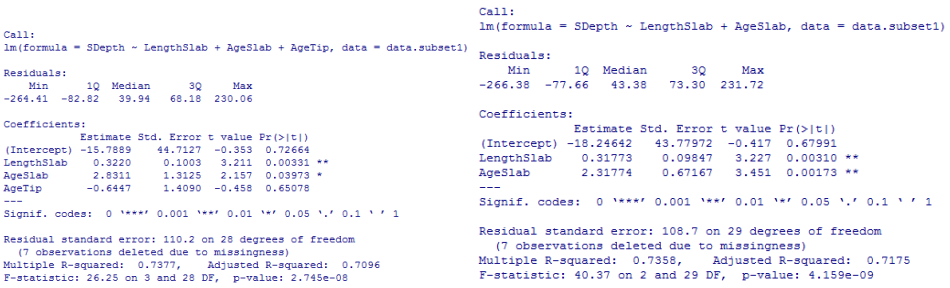
Figure 5.13: Steps of modelling the Slab Length using Multiple Linear Regression with Minister & Jordan's velocity model

The predictor V_c in this model is calculated using Chase's velocity model, so the model should also be analysed and expressed using Minister & Jordan's velocity model. Figure 5.13 shows the steps for calculating the new model which are the same like before. SDepth is shown to be insignificant in this case as well. What changes now is that the R-squared values are slightly increased to 72 % and the coefficients that are also slightly altered as shown in the model:

$$LengthSlab = -25.3 + 4.5 AgeSlab + 51.3 V_c$$

- Maximum Depth of Benioff Zone - SDepth Model

  The SDepth parameter is highly correlated with the variables LengthSlab, AgeSlab and AgeTip according to the correlations table and they will be used as predictors in the first model as shown in figure 5.14a. The results of the model indicate that the AgeTip is insignificant variable as it has high p-value. After excluding AgeTip from the predictors, a new model is created and evaluated as shown in figure 5.14b. Both of the LengthSlab and the AgeSlab are significant as their p-values are low and the overall model has high R squared



(a) First assessed model of M.L.R. with three explanatory parameters

(b) Final assessed model of M.L.R. with two explanatory parameters

Figure 5.14: Steps of modelling the Maximum Depth of Benioff Zone using Multiple Linear Regression

value which proximately equals to 72. The final model of SDepth is:

$$SDepth = -18.25 + 2.32AgeSlab + 0.32LengthSlab$$

5.3.4 Model Diagnostics

Applying linear regression, automatically means that some assumptions are respected. These assumptions are that the errors (residuals) of the models are normally distributed, that they have the regression line as a centre and that their variance does not change as a function of x. It is important then to perform some investigations and evaluations of our models using model diagnostics. The previously expressed models will be investigated through the regression diagnostic plots of residuals versus fitted (predicted) values, normal quantile-quantile plot of standardized residuals, scale-location and Residuals versus leverage.

- Residuals versus fitted values plot

  Part of the dependent variable are the residuals. They are points that the model could not explain, so they are the best available sample of errors of the regression model. A residual is given after subtracting a predicted value from the actual value of the dependent variable and expressed as the vertical distance between a point and the regression line. Under the assumptions of a linear regression model, the variability of residuals over the range of the dependent variable (fitted values) should not form any pattern on this plot.

  Looking at the residuals vs fitted diagnostic plots of the figures 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23 it is easy to observe that the different models give completely different results. The most positive results are given mainly from the model of DipD but also from the models of DipI and Length Slab. The red lines are smooth curves that pass through the actual residuals, they are relatively flat and lie close to the gray dashed lines ($y = 0$). This is the result that we hoped to observe on this plots. On the other hand, the curves related to the

models of M, Dd, SDepth and Gap (a-t) seem to form some curves which escape from the grey line indicating that a linear model may not be a good fit to these data.

- Normal quantile-quantile plot of standardized residuals plot

  As mentioned in the assumptions of linear regression, the residuals of the models must follow a normal distribution and this plot evaluates that assumption. That means that a well expressed model should follow a straight line ($y = x$) on the normal quantile-quantile plot. Deviations from this line may imply that the errors fail to follow a normal distribution. Some small deviations though, should be expected, particularly near the ends.

  After evaluating the results of the normal Q-Q plots in figures 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23, the models can be divided in three groups. Firstly the group with the best results includes the models of DipI and Length Slab. Almost all their points lie exactly on the $y = x$ dashed line with the exception of a few small deviations. The second group includes DipD, SDepth and M models which give satisfactory results with only some small deviations focused mostly near the ends. Finally the less successful group which includes curved patterns on the plots contains Dd and Gap (a-t) models.

- Scale-location plot

  This plot is very similar to the first one, with the difference of using the square root of the standardized residuals so that they have a mean of zero and a variance of one. In this plot the sign on the residual is eliminated and the small residuals are plotted on the bottom while the large ones on the top. Like before, the plot should not form any discernible pattern.

  Looking at the scale-location plots in figures 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23, red lines that are relatively flat are expected to be observed as they depict the trends. The models of Length Slab, Dd, DipI and DipD give such lines while M, SDepth and Gap models seem to define a trend confirming the results of the residuals vs fitted diagnostic plots.

- Residuals versus leverage plot

  Leverage consists a statistical measure that expresses how much each data point influences the regression. Leverage reflects both the distance of a point from the centroid and its isolation. Points that lie far from the centroid have greater leverage, and this leverage increases the fewer the points that are nearby. That happens because the regression line must always pass through the centroid. This plot contours values of Cook's distance as well, which is a statistic that tries to identify the points that have more influence compared to the others and measures how much would the regression change if a point was excluded. High leverage and large residuals increase

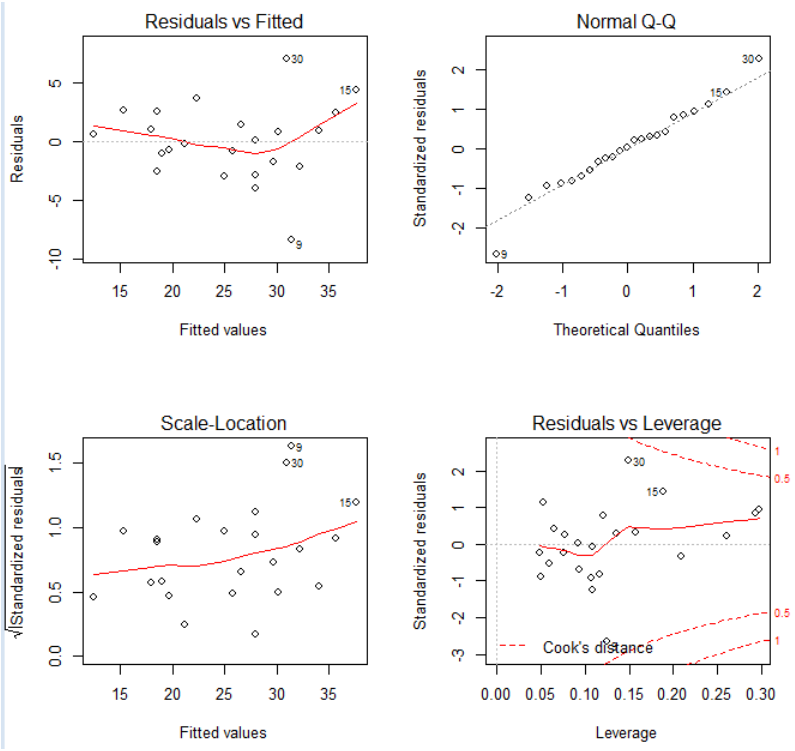


Figure 5.15: Intermediate Dip Model Diagnostic Plots (Chase's velocity model)

Cook's distance, so a point with large residual and far from the centroid can distort the regression immensely. That is why the points on this plot should have short Cook's distance (less than 0.5).

Checking the bottom right plots in figures 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23, it is easy to notice that the best results are given from the DipD model as the red line is smooth and stays close to the horizontal dashed line while no points have large Cook's distance. The next most positive results are given from the DipI and Length Slab models although the later has one point having Cook's distance little more than 0.5. The models of M and Dd, although having all their points with short Cook's distances, seem to have curved red lines and finally SDepth and Gap (a-t) models which have both curved red lines and one point with Cook's distance little longer than 0.5.

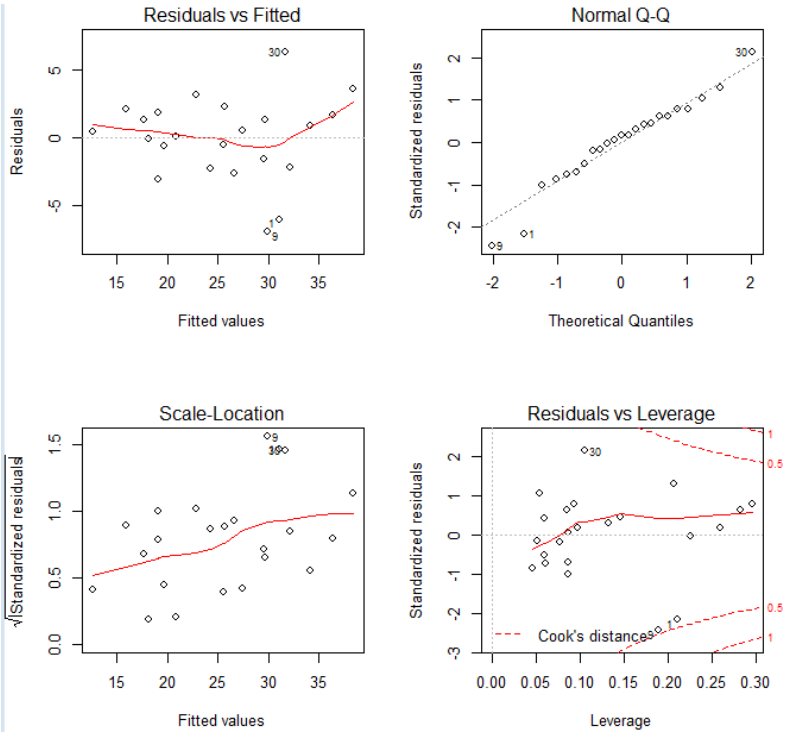


Figure 5.16: Intermediate Dip Model Diagnostic Plots (Minister & Jordan's velocity model)

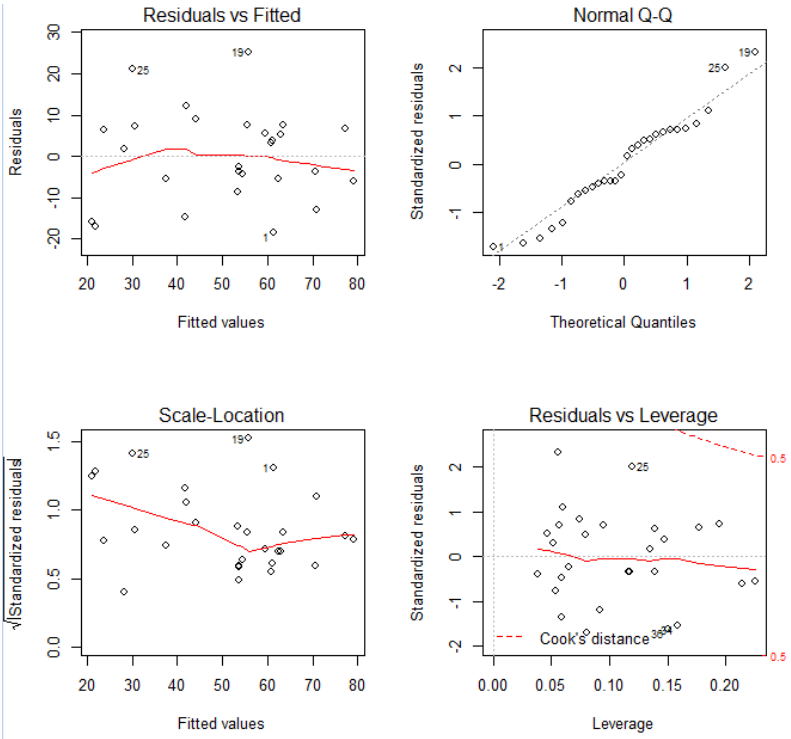


Figure 5.17: Deep Dip Model Diagnostic Plots

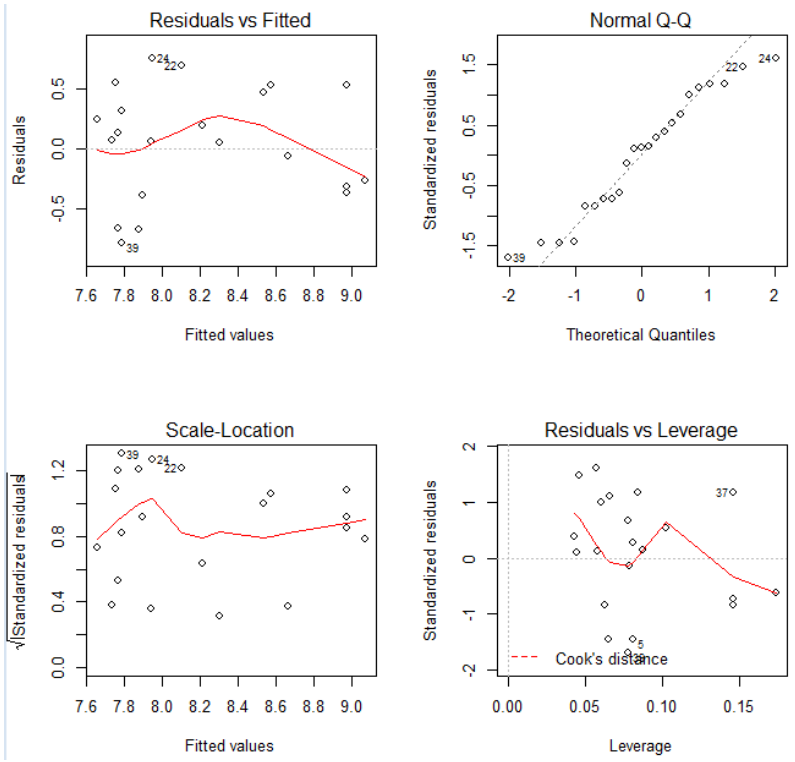


Figure 5.18: Maximum Cumulative Earthquake Moment Model Diagnostic Plots

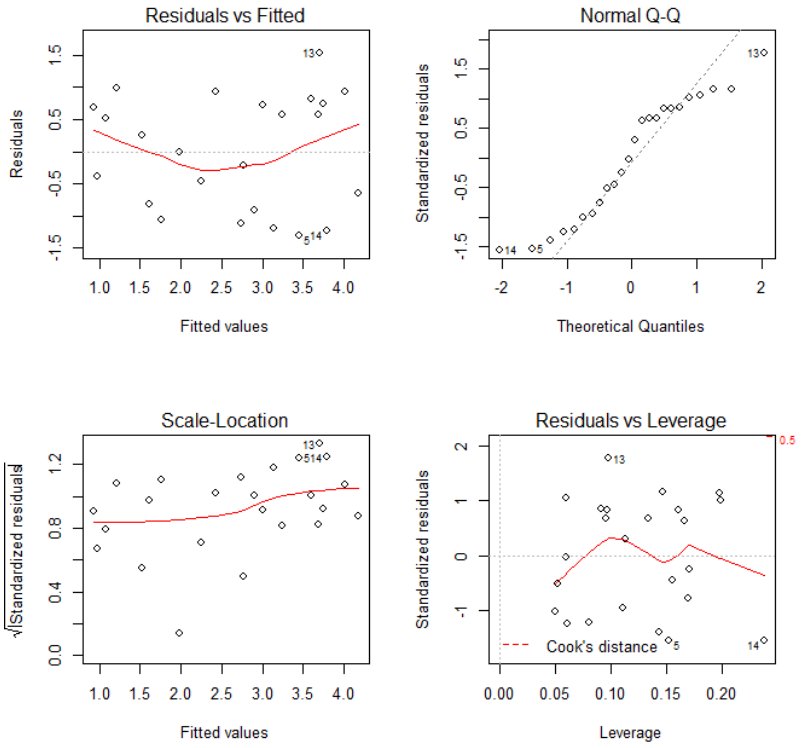


Figure 5.19: Relative Trench Depth Model Diagnostic Plots [59]

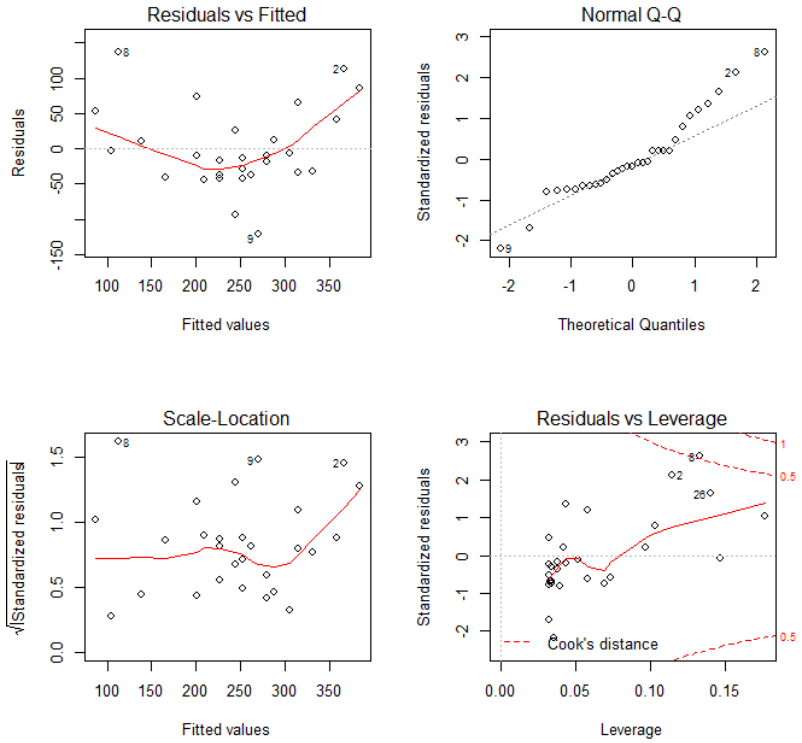


Figure 5.20: Gap Between Arc & Trench Model Diagnostic Plots

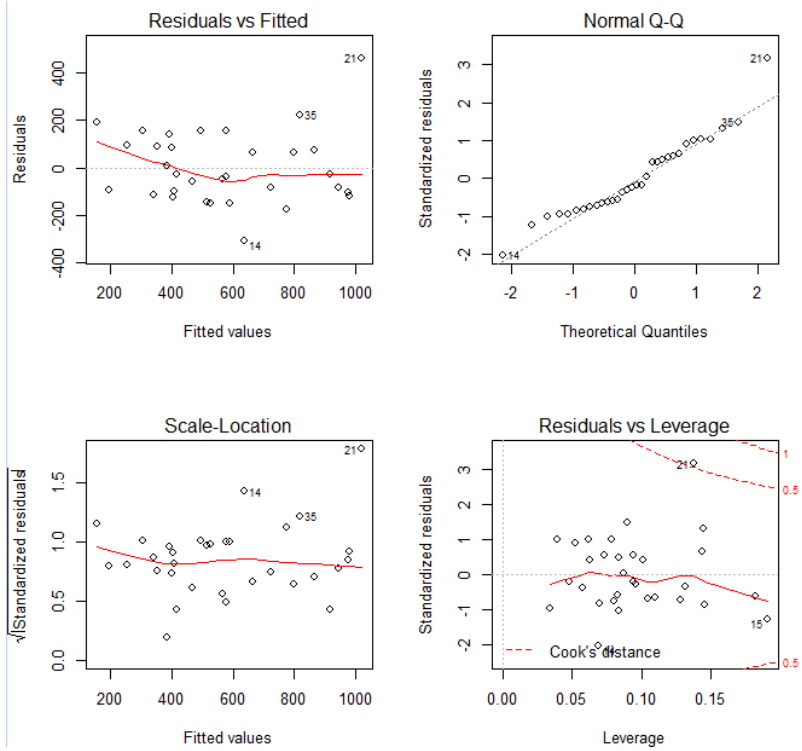


Figure 5.21: Length of Slab Model Diagnostic Plots (Chase's velocity model)

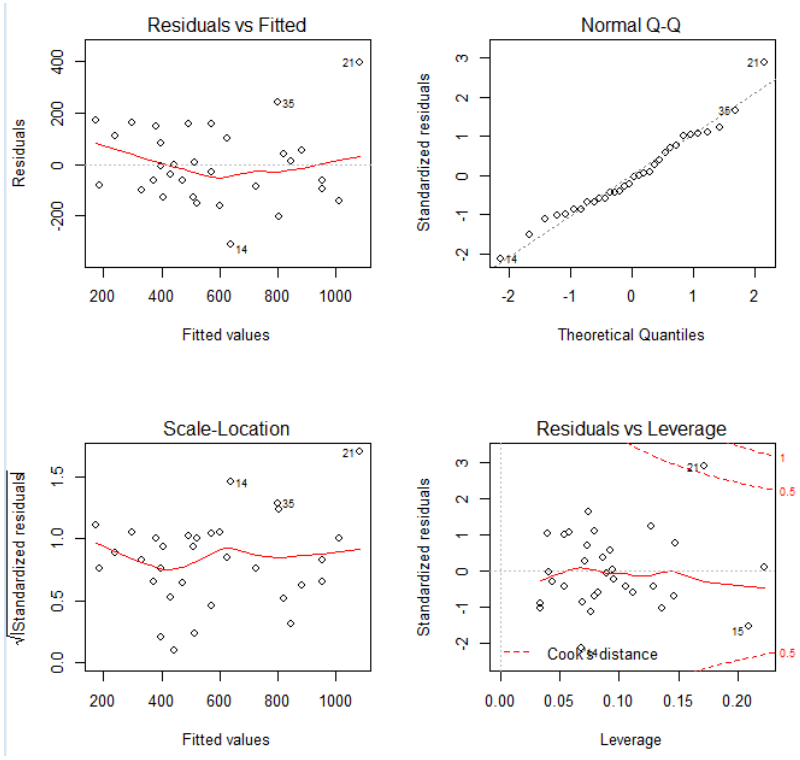


Figure 5.22: Length of Benioff Zone on the Slab Model Diagnostic Plots (Minister & Jordan's velocity model)

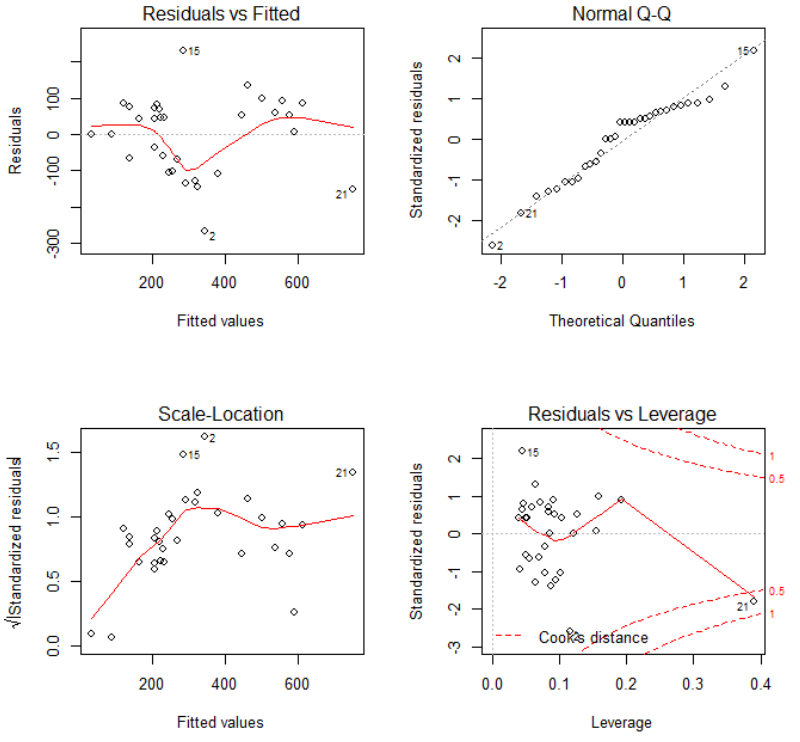


Figure 5.23: Maximum Depth of Benioff Zone Model Diagnostic Plots

5.3.5 Results

To finally evaluate the nine models that were built in the previous paragraphs we combine the results of the R squared values of each model and the feedback from the regression diagnostic plots that were analysed previously. Table 5.3 contains the Multiple and the Adjusted R-squared values of each model as well as evaluation score for each diagnostic plot. The meaning of the evaluation score is as follows:

- 1 = very good results
- 2 = satisfying results
- 3 = not very good results

The models that are proved to be well defined, describing successfully the dependent variables, are those of Intermediate Dip (the model that used Minister & Jordan's velocity data gave slightly better results), Deep Dip, Slab Length (once again the model that used Minister & Jordan's velocity data gave slightly better results) and SDepth (although having high R squared values, plots seem to be problematic only because of a couple of outliers). The models that do not give very satisfying results are these of Maximum Cumulative Earthquake Moment, Relative Trench Depth and Gap between Arc & Trench. It is worth mentioning that their results are not poor but comparing them with the rest of the models, they fail to be adequately linearly expressed by their explanatory variables.

Model	R^2	**Adjusted** R^2	**Residuals vs Fitted**	**Normal Q-Q**	**Scale - Location**	**Residuals vs Leverage**
$DipI(C) = 36.64 - 0.07Gap(a-t) + 0.082DipU$	82%	80%	2	1	1	2
$DipI(M) = 36 - 0.07Gap(a-t) + 0.097DipU$	84%	83%	2	1	1	2
$DipD = 40 + 0.9DipI - 0.14AgeArc$	70%	68%	1	2	2	1
$M = 7.6 + 0.01AgeArc$	53%	51%	3	2	3	3
$Dd = -0.83 + 0.08DipI + 0.02AgeSlab$	59%	55%	3	3	1	3
$Gap(a-t) = 472.6 - 8.8DipI$	66%	65%	3	3	3	3
$LengthSlab(C) = 4 + 4.3AgeSlab + 46V_c$	72%	70%	2	1	1	2
$LengthSlab(M) = -25.3 + 4.5AgeSlab + 51.3V_c$	74%	72%	2	1	1	2
$SDepth = -18.25 + 2.32AgeSlab + 0.32LengthSlab$	74%	72 %	3	2	3	3

Table 5.3: Models' evaluation table

Chapter 6

Reconstruction and Visualization

6.1 Plate Tectonics Reconstruction

In part I, subduction theory is analysed, while in chapter 5 mathematical and statistical methods are used to model different subduction parameters. Theoretical knowledge and different models that have been created, give insight and understanding of the subduction mechanism. The next step is the development of visualisation tools which give a better and more spherical understanding of the interactions between tectonic plates and different geological phenomena related to it including subduction.

Plate-tectonic reconstruction and its visualisation, are useful tools in a variety of contexts like geophysics, research in geology, exploration of hydrocarbon and mineral resources, paleotopography, climate change modelling etc. By plate-tectonic reconstruction is meant the calculation of the positions and the orientation of tectonic plates relative to each other or to other reference frames - like Earth's magnetic field or hotspot groups - at a specific instant in Earth's history. The visualisation of such reconstructions helps the understanding of the processes that occur on the Earth's surface and near the subsurface by giving realistic models of the evolution of the Earth's plate tectonics. In order for a researcher to be able to trace plate motions through time, both geological and geophysical features are usually embedded in the simulation of plate reconstruction.

The modern software used for plate reconstruction visualisation, usually makes use of time-sequence reconstructions that are used to animate the plates' motions by producing time-derivative information like kinematic. Reconstructing continents is an important feature of software which facilitate the study and exploration of the distant geological past. In addition it must incorporate, edit and reconstruct all geological, geophysical and paleogeographic data within a plate tectonic environment. Finally visualisation should be expressed through real-time graphical manipulation of both tectonic plate motion and plate reconstruction.

6.2 Common Software for Visualizing Reconstruction

- **Plates**

  Plates is a desktop software for research into plate tectonics and geologic reconstruction. The objective is to construct accurate and high resolution models of the plate movement in the past and the present. It is a software developed mainly at the University of Texas Institute for Geophysics with the help of external collaborations. It is no longer in use and it consists one of the first complete attempts of creating such a software which allowed the reconstruction of data like coastal outlines, by using geographical points. Compared with modern products it lacks in resolution and functionality. [42]

- **SPlates**

  SPlates is a computer software developed later than Plates, which was designed to restore different geological data to their initial positions on Earth in order to produce paleogeographic maps. It provides analytical tools which help to determine changes concerning the rates and the directions of plate movement on Earth through time. This helps determine the tectonic forces that occur within the plates' margins as they move respectively to each other. The SPlates project was funded by StatoilHydro which is a Norwegian company that deals with hydrocarbon exploration and is no longer in use as it is replaced by other more up to date and advanced programs. [55]

- **PPlates**

  PPlates is a desktop software for tectonic reconstruction which allows tectonics research and structural geology. This program is different and more evolved compared to the previous ones. PPlates does not consider plates to be rigid and does not simply simulate their movements. It also models and visualises the deformation of the lithosphere by expressing it in terms of deformable and tearable 3D meshes. In previous software, properties like Euler pole and rotation data were the primary input while in this case they have become derivative properties. The mesh here plays the role of an in-built coordinate system and gives the ability to the crustal deformations to conserve properties like mass or volume, isostasy and rheology. PPlates was developed by the Research School of Earth Sciences in the Australian National University. [63], [43]

- **GPlates**

  GPlates is a new generation open-source desktop software for the calculation and visualisation of plate-tectonic reconstructions. Its functionality is similar to GIS (Geological Information System) with the added dimension of geological time. The motivation for the development of GPlates among others, was the visualisation of reconstructed geological, geophysical and paleogeographic data in many formats, the link of plate kinematics with geodynamic models

and the quick production of high - quality paleographic maps. It offers all the previously familiar features with the addition of many improvements. The multiple frame-per-second animations results in smoother visualisation of reconstructions, the raster data are seamlessly displayed on the surface of the globe, the reconstructions can be manipulated interactively with constant real-time feedback and all these are given through a modern and user friendly graphical interface. GPlates is developed by a group of scientists and program developers from the EarthByte Project in the School of Geosciences at the University of Sydney, the Division of Geological and Planetary Sciences (GPS) at Caltech and finally the Centre of Geodynamics at the Norwegian Geological Survey (NGU). [5], [70]

- **4DPlates**

4DPlates is one of the latest and most modern software applications which was developed to reconstruct the positions of geological data in the geologic past and later display them in high resolution. Compared to the previous software, 4DPlates can support larger and higher precision data sets which describe both the surface and the interior structure of the Earth. The visualisation results are very smooth as it achieves around 30 to 50 frames per second and that makes it suitable for quick and detailed reconstruction, display and manipulation of high resolution geographic data sets. The user can choose to display multiple layers of data simultaneously, by changing the transparency of each level so that he can notice the changes in each one and can also generate video in different formats which is a new component in such software. The graphics processing unit (GPU) of 4DPlates is one of its most powerful aspects as the user can perform usual actions like data-masking, calculations and spherical rotations but he can also implement interactive raster colour mapping, surface lighting and scaling. These visual manipulation of data can reveal hidden details in geological data. Finally 4DPlates provides the user with some useful and new functionalities which make the environment friendlier and extraction of the results easier. The user can customise the gridded data sets by applying algebraic functions that he defines, he can also compare side by side two different models at different geological times and finally he can place flow-lines to illustrate the relative motion of the plates. 4DPlates was funded by Statoil and was developed in collaboration by Kalkulo and the department of Computational Geoscience in Simula Research Laboratory. [7]

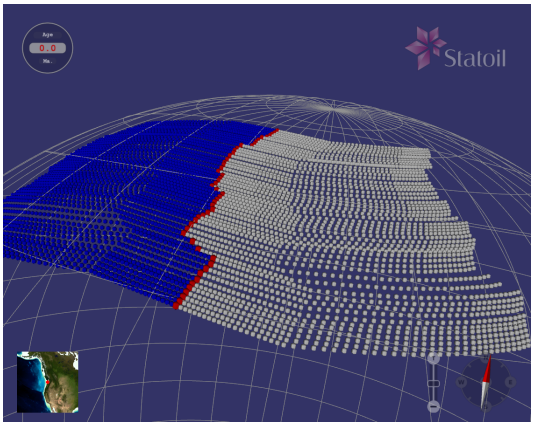
Figure 6.1: Location of the Cascadia subduction zone [61]

6.3 My Work

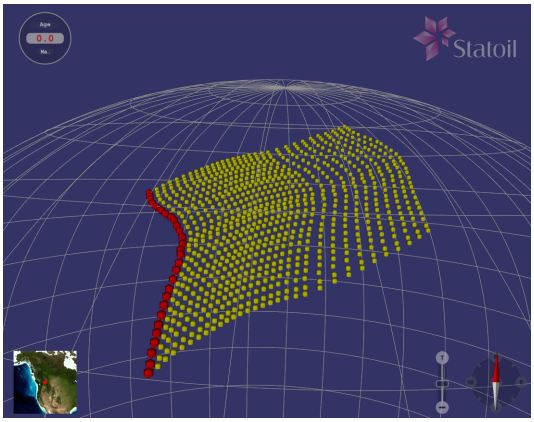
In this paragraph the reconstruction of the subducting Juan de Fuca plate at the Cascadia subduction zone for the last 30 Ma is performed. The reconstruction and its visualisation is achieved by applying some of the models of the section 5.3.3 in the 4DPlates software, combined with the functionalities that the later provides.

More specifically, the bathymetrical reconstruction of the current Benioff zone on the Juan de Fuca plate is reconstructed and visualized. For this purpose the model which is mainly used is the one of the Maximum depth of the Benioff zone on the slab (SDepth) and is described in the previous paragraph.

The Cascadia subduction zone is a convergent plate boundary between the North America and Juan de Fuca plates and stretches from Northern Vancouver Island to Cape Mendocino in northern California as shown in figure 6.1. Juan de Fuca is located between the Pacific plate and the North America one, and was formed during the Oligocene (from about 33.9 million to 23 million years before the present time) after the Farallon plate broke into a series of smaller plates. [15]



(a) Flowlines showing the spreading of lithosphere between the Juan de Fuca plate and Pacific one.



(b) Flowlines showing the path of motion of the Juan de Fuca plate as it subducts under the North America plate.

Figure 6.2: Flowlines which depict the paths of motion of the Juan de Fuca plate. Red dots indicate present day location of spreading or subduction.

6.3.1 Process of reconstruction

The work done for the reconstruction and visualization of the Cascadia subduction zone can be summarized in the following steps.

- 4DPlates gives the possibility to the user to study every area on the earth he may be interested in. For this reason, the first step was to zoom in the Cascadia area and mark the zone of interest for the project. This was done by using some functionalities of the 4DPlates. Firstly using flowlines at the tectonic boundaries of the Juan de Fuca plate (North America plate and Pacific plate) the paths of spread (fig. 6.2a) and subduction (fig. 6.2b) of the plate were visualized. Using these paths, we drew a polygon which enclosed all the area that the Juan de Fuca plate covered while subducting the last 30 Ma .
- The next step was to create a grid file with both the ages of the

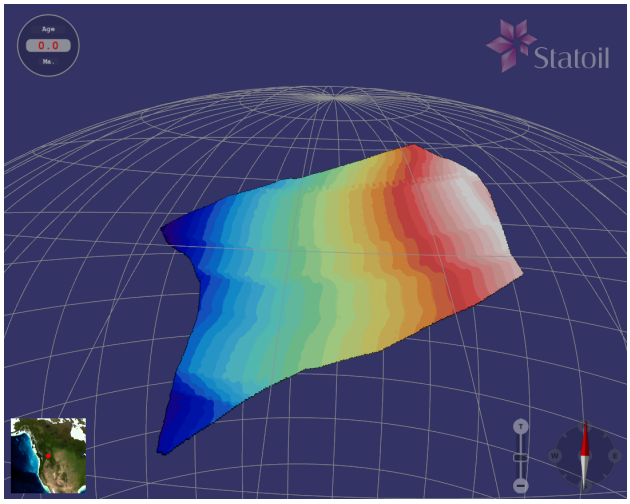


Figure 6.3: Age of the Lithosphere

lithosphere and subduction in the polygon marked before. The age of the slab (or lithosphere) is the most important parameter which defines its depth and the age of subduction is useful to distinguish if a part of the slab is subducted or not.

The first grid file, containing the lithospheric ages, was created in two different ways for reasons of research curiosity giving similar results. The first one was using palaeo-age data downloaded from EarthByte [1]. Combining this data with the polygon drawn before, a grid file of the ages concerning only the coordinates of the slab was created. The second way of creating the age grid file which was also the one finally used, was using the functionalities of 4DPlates and the age masking that it contains. As seen in figure 6.3, its visualization gives the age of the part of the lithosphere selected. In this figure, the ages denoted with white colour are older than the ones denoted with blue. This was the expected result as the ridge where new oceanic lithosphere is created is on the western side of the Juan de Fuca plate on the divergent boundary with the Pacific plate.

The second grid file, containing the age of subduction of the lithosphere, was created using again the functionalities of 4DPlates. The result can be seen in figure 6.4. In this figure, white colour denotes 30 Ma age while dark blue denotes present day or else 0 Ma. As expected, the white colour is on the eastern part of the polygon as these areas where the first to be subducted at 30 Ma.

- The last step was to reconstruct and visualize the maximum depth of the Benioff zone of the subducting slab. This was done once again by using one of the functionalities of 4DPlates, to apply short algorithms and mathematical formulas over a grid layer. The formula used in this case is the result of the interpolation between two equations. The

[1] http://earthbyte.org/Resources/palaeoagegrid2008.html

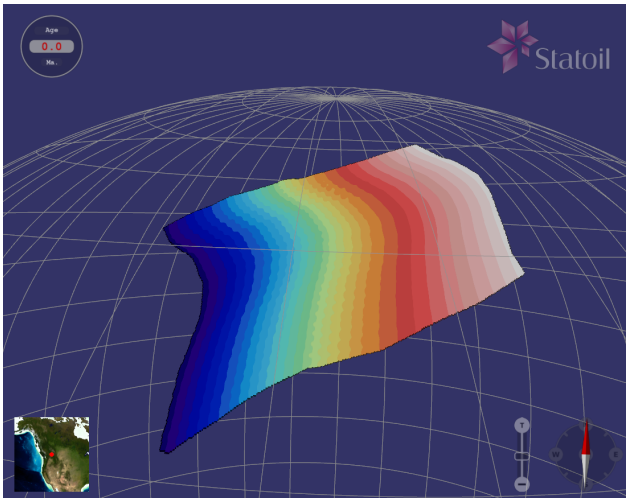


Figure 6.4: Age of Subduction

first one

$$x = 2600 + 345\sqrt{Ageof Lithosphere}$$

explains the sea-floor topography of the oceanic lithosphere before it subducts and determines the ocean depth [66]. The second equation

$$x = 2.32 * Ageof Lithosphere + 0.32 * LengthSlab - 18.25$$

is the one derived from the multiple linear regression which was applied on subduction parameters in a previous paragraph (5.3.3) and describes the depth of the slab after it was subducted. Interpolation between the two equations was performed in order to achieve a smoother result when visualizing the slab. The result is as follows:

```
if (gt >=AoS+window/2)
{
x=2600+345 \sqrt{AoL}
}
else if (gt>=AoS-window/2)
{
x=(0.5-AoL/window)*(2.32*AoL + 0.32*LengthSlab-18.25)
+(0.5+AoL/window)*(2600+345 \sqrt{AoL})
}
else
{
x= 2.32*AoL + 0.32*LengthSlab-18.25
}
```

where
AoS: the Age os Subduction
AoL: the Age of Lithosphere as described before
gt: the geological time

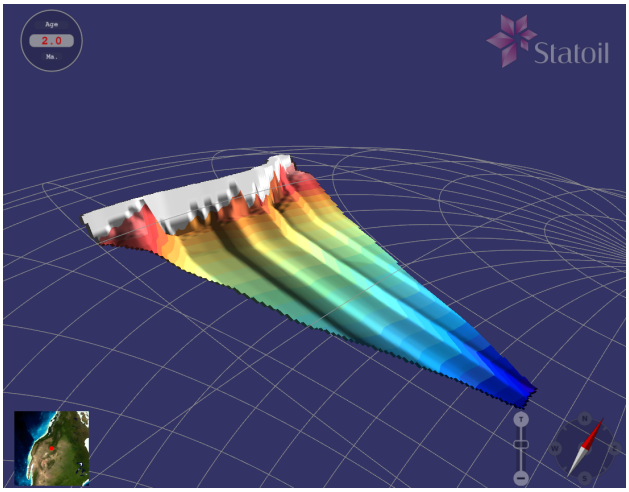
6.3.2 Results

The objective of the previous work flow was to visualize the subducting Juan de Fuca plate and try to peek under the North America plate in order to see how the existing slab subducted through the last 30 Ma.

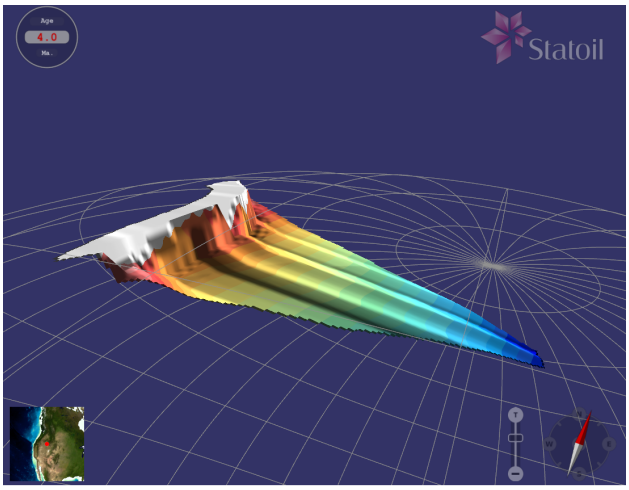
The result can be seen in figure 6.5 where the slab is depicted in different time instances. The colours of the slab denote its depth and indicate a downward inclination as expected. The dark blue coloured areas on the slab correspond to deep depths while the white colours correspond to relatively swallower depths. The time instances in figure 6.5 are at 2 Ma, 4 Ma and 14 Ma and demonstrate the evolution of the subducting slab. At 2 and 4 Ma the largest part of the slab is subducted and only a small part of the slab is unrevealed. At 14 Ma it is obvious that the largest part of the slab is not yet subducted - white colour - while the subducted one is shorter and in shallower depth.

Observing the reconstructed slab, some depth anomalies can be seen running on the horizontal axe. These anomalies are actually transform faults around which a slab may have sharp age variations. These variations cause depth anomalies on the slab which are also obvious in figure 6.5.

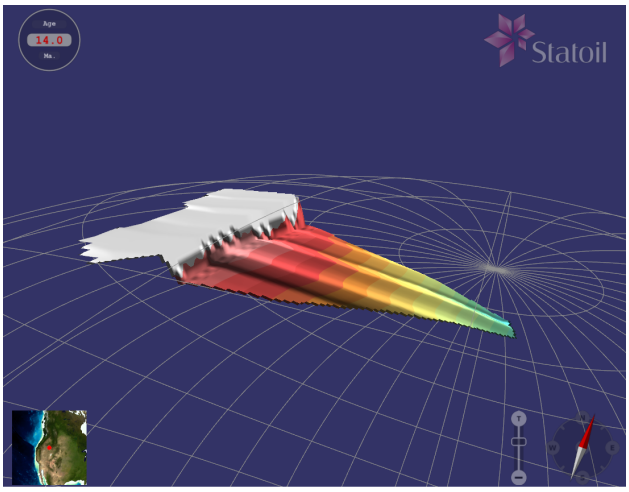
Finally in figure 6.6 the Juan de Fuca subducting slab is visualized relative to the overriding North America plate so that it becomes clearer how the subduction evolved through time and give the general picture of it. The first instance is at 2 Ma where only a small proportion of the slab is revealed while the rest of it is under the North America plate going down in big depth. The second instance, at 13 Ma demonstrates that a big part of the subducting plate is not yet subducted while the slab under the American plate is found in shallower than before depths. The final instance depicts the subduction at 23 Ma, when the largest part of the Juan de Fuca plate is not subducted and only a small part of the plate lays under the America one in even shallower depths than before.



(a) Instance of the slab reconstruction at 2 Ma.

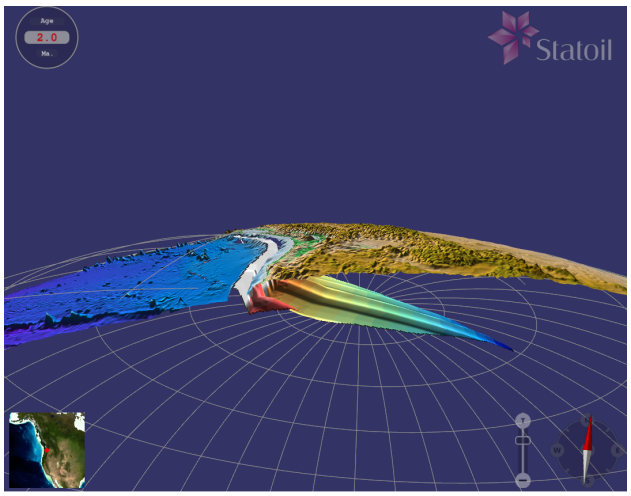


(b) Instance of the slab reconstruction at 4 Ma.

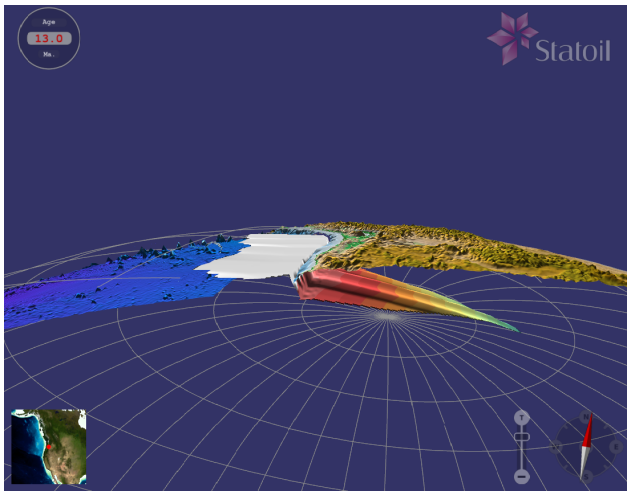


(c) Instance of the slab reconstruction at 14 Ma.

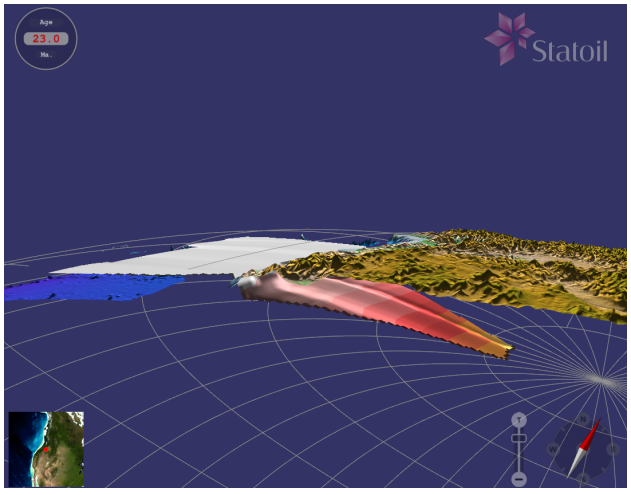
Figure 6.5: Different time instances of the Juan de Fuca slab reconstruction.



(a) Instance of the Cascadia's reconstruction at 2 Ma.



(b) Instance of the Cascadia's reconstruction at 13 Ma.



(c) Instance of the Cascadia's reconstruction at 23 Ma.

Figure 6.6: Different time instances of the Juan de Fuca reconstruction which include the North American plate under which the slab subducts.

Part III

Conclusion

Chapter 7

Results

Using the R software environment for statistical computing, correlation and clustering was applied on 20 different subduction parameters in order to detect possible relationships within them. These relationships were later modelled with the help of multiple linear regression, giving the following results.

- The intermediate dip (DipI) of the subducting slab can be defined and predicted by the arc-trench gap (Gap (a-t)) and the descent angle of slab into the mantle (DipU). The models which were derived are:

  using Chase's model for DipU:
  $DipI = 36.64 - 0.07Gap(a-t) + 0.082DipU$

  using Minister and Jordan's model for DipU:
  $DipI = 36 - 0.07Gap(a-t) + 0.097DipU$

  The model diagnostics showed that these two models can predict the intermediate dip with good results.

- The deep dip (DipD) of the subducting slab proved to be explained by the intermediate dip (DipI) of the subducting slab and the duration of the subduction (AgeArc) in that zone. The resulted model is:

  $DipD = 40 + 0.9DipI - 0.14AgeArc$

  The model diagnostics in this case showed that the deep dip is described very well by this model.

- The maximum cumulative earthquake moment (M) of a subduction zone can only be described by the duration of the subduction (AgeArc) in the area. The model which is derived is:

  $M = 7.6 + 0.01AgeArc$

  According to the model diagnostics, the maximum cumulative earthquake moment is described in a sufficient way by its model.

- The relative trench depth (Dd) of a subduction zone can be explained by both the intermediate dip (DipI) and the age of the slab at the trench (AgeSlab). The model of Dd is as follows:

  $Dd = -0.83 + 0.08DipI + 0.02AgeSlab$

  Relative trench depth is also sufficiently described by its model according to the model diagnostics.

- The arc-trench gap (Gap (a-t)) can only be described by the intermediate dip (DipI) of the subducting slab. The resulting model is:

  $Gap(a-t) = 472.6 - 8.8DipI$

  The model diagnostics proved that the expression of the arc-trench gap by the intermediate dip is sufficient.

- The maximum depth of the Benioff zone on the slab (SDepth) can be predicted by using the length of the slab (LengthSlab) and its age (AgeSlab). The model which resulted is :

  $SDepth = -18.25 + 2.32AgeSlab + 0.32LengthSlab$

  According to the model diagnostics the maximum depth of the Benioff zone on the slab can be described in a good way from the LengthSlab and AgeSlab variables.

- The last subduction parameter that was modelled is the length of the subducting slab (Length Slab) which can be explained by the age of the slab at the trench (AgeSlab) and the convergence rate V_c of the subduction zone. The derived models are:

  using Chase's velocity model for DipU :
  $LengthSlab(C) = 4 + 4.3AgeSlab + 46V_c$

  using Minister and Jordan's velocity model for DipU:
  $LengthSlab(M) = -25.3 + 4.5AgeSlab + 51.3V_c$:

  The model diagnostics proved that the final two models explain the length of the subducting slab in a good way.

Using the software for plate reconstruction and visualisation called 4DPlates, the model of maximum depth of the Benioff zone on the slab was applied on the Cascadia subduction zone. The age and bathymetry of the Juan de Fuca plate were reconstructed so that the evolution of the subduction during the last 30 Ma could be visualized.

The visual result demonstrates the form of the current subducted slab throughout the last 30 Ma and even depicts the transform faults on the Juan de Fuca plate which cause depth anomalies on the slab.

Chapter 8

Future Work

The project has been successful in creating statistical models within subduction parameters and later reconstructing and visualizing a subducting plate called Juan de Fuca at the Cascadia subduction zone. There are though some proposals about future work that could be considered in order to possibly improve the current results.

- The subduction parameters that were examined and later modelled, were taken from Richard D. Jarrard's work done in 1986. It is true that this kind of data do not get old since geological changes occur within million of years. Jarrard's work has also been considered one of the best resources for data concerning subductions and it has been widely used for later and further research. However, in the future, it would be interesting for different and more contemporary data to be used. Geoscience is a discipline that constantly evolves and some values could be changed by applying new techniques. These data may be fragmented and much more difficult to be assembled but it is worth to be gathered and processed so that new models are created following the steps made in this project. The new models can be later compared and assessed with the current ones.

- The subduction parameters in the current project, were linearly modelled using multiple linear regression. This resulted in some very good and explanatory models while a few of them gave only satisfying results. In future work, these models could be improved by not only dealing with linearity which could prove to give more descriptive results.

- The visualization part of this project is a novel work which makes the result unique and provides incentive for further research. At this point the reconstruction and visualization of the slab is a result of statistical modelling among parameters of many subduction zones. It would be interesting though to gather existing data concerning the Cascadia zone exclusively and reconstruct the slab following the new results. This could give details about the evolution of the subduction in this zone which are now hidden. It would be also interesting

to model other subduction parameters as well, whose visualization would give more information about the subducting slab.

Bibliography

[1] Don L Anderson and Adam M Dziewonski. 'Seismic tomography'. In: *Scientific American* 251 (1984), pp. 60–68.

[2] Brian F Atwater et al. 'Surviving a tsunami: Lessons from Chile, Hawaii, and Japan'. In: *US Geological Survey Circular* 1187 (1999), pp. 1–18.

[3] Markus Bêath. *Introduction to seismology*. Birkhauser, 1979.

[4] Peter Bird. 'An updated digital model of plate boundaries'. In: *Geochemistry, Geophysics, Geosystems* 4.3 (2003).

[5] James A Boyden et al. 'Next-generation plate-tectonic reconstructions using GPlates'. In: *Geoinformatics: cyberinfrastructure for the solid earth sciences* (2011), pp. 95–114.

[6] Clement G Chase. 'Plate kinematics: the Americas, East Africa, and the rest of the world'. In: *Earth and planetary science letters* 37.3 (1978), pp. 355–368.

[7] Stuart R Clark et al. '4DPlates: On the fly visualization of multilayer geoscientific datasets in a plate tectonic environment'. In: *Computers & Geosciences* 45 (2012), pp. 46–51.

[8] Kent C Condie. *Plate tectonics & crustal evolution*. Elsevier, 2013.

[9] *Convergent and Transform Boundaries*. July 2009. URL: http://cosscience1.pbworks.com/w/page/8286031/Lesson%2010-5%20Convergent%20and%20Transform%20Boundaries.

[10] Timothy A Cross and Rex H Pilger. 'Controls of subduction geometry, location of magmatic arcs, and tectonics of arc and back-arc regions'. In: *Geological Society of America Bulletin* 93.6 (1982), pp. 545–562.

[11] Adam M Dziewonski and Don L Anderson. 'Preliminary reference Earth model'. In: *Physics of the earth and planetary interiors* 25.4 (1981), pp. 297–356.

[12] FC Frank. 'Curvature of island arcs'. In: (1968).

[13] Eric L Geist et al. *Helping Coastal Communities at Risk from Tsunamis, the Role of US Geological Survey Research*. US Department of the Interior, US Geological Survey, 2000.

[14] *Geologic Provinces of the United States: Rocky Mountains*. Oct. 2014. URL: http://geomaps.wr.usgs.gov/parks/province/rockymtn.html.

[15] Rob Govers and Paul Th Meijer. 'On the dynamics of the Juan de Fuca plate'. In: *Earth and Planetary Science Letters* 189.3 (2001), pp. 115–131.

[16] Marc Grunberg, Stéphane Genaud and Catherine Mongenet. 'Seismic ray-tracing and Earth mesh modeling on various parallel architectures'. In: *The Journal of Supercomputing* 29.1 (2004), pp. 27–44.

[17] Michael Gurnis, Chad Hall and Luc Lavier. 'Evolving force balance during incipient subduction'. In: *Geochemistry, Geophysics, Geosystems* 5.7 (2004).

[18] Bradford H Hager and Richard J O'connell. 'Subduction zone dip angles and flow driven by plate motion'. In: *Tectonophysics* 50.2 (1978), pp. 111–133.

[19] Chad E Hall et al. 'Catastrophic initiation of subduction following forced convergence across fracture zones'. In: *Earth and Planetary Science Letters* 212.1 (2003), pp. 15–30.

[20] Arnauld Heuret et al. 'Physical characteristics of subduction interface type seismogenic zones revisited'. In: *Geochemistry, Geophysics, Geosystems* 12.1 (2011).

[21] RD Hyndman and K Wang. 'The rupture zone of Cascadia great earthquakes from current deformation and the thermal regime'. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 100.B11 (1995), pp. 22133–22154.

[22] *Introduction to Teaching Plate Tectonics*. URL: http://geology.com/nsta/.

[23] Richard D Jarrard. 'Relations among subduction parameters'. In: *Reviews of Geophysics* 24.2 (1986), pp. 217–284.

[24] Thomas H Jordan. 'Structural geology of the Earth's interior'. In: *Proceedings of the National Academy of Sciences* 76.9 (1979), pp. 4192–4200.

[25] Philip Kearey, Keith A Klepeis and Frederick J Vine. *Global tectonics*. John Wiley & Sons, 2013.

[26] Philip Kearey, Keith A Klepeis and Frederick J Vine. *Global tectonics*. John Wiley & Sons, 2013.

[27] Serge Lallemand, Arnauld Heuret and David Boutelier. 'On the relationships between slab dip, back-arc stress, upper plate absolute motion, and crustal nature in subduction zones'. In: *Geochemistry, Geophysics, Geosystems* 6.9 (2005).

[28] Xavier Le Pichon, Jean Francheteau and Jean Bonnin. *Plate tectonics*. Elsevier, 2013.

[29] Harold L Levin. *The earth through time*. John Wiley & Sons, 2009.

[30] Andrea L Llenos and Jeffrey J McGuire. 'Influence of fore-arc structure on the extent of great subduction zone earthquakes'. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 112.B9 (2007).

[31] William F McDonough and S-S Sun. 'The composition of the Earth'. In: *Chemical geology* 120.3 (1995), pp. 223–253.

[32] J Bernard Minster and Thomas H Jordan. 'Present-day plate motions'. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 83.B11 (1978), pp. 5331–5354.

[33] RD Müller and TCW Landgrebe. 'The link between great earthquakes and the subduction of oceanic fracture zones'. In: *Solid Earth* 3.2 (2012), pp. 447–465.

[34] Alan E Mussett and M Aftab Khan. *Looking into the earth: an introduction to geological geophysics*. Cambridge University Press, 2000.

[35] *NF-GEO 3310/4310: Imaging in seismology*. Nov. 2007. URL: http://www.uio.no/studier/emner/matnat/ifi/INF-GEO4310/h11/undervisningsmateriale/seismology.pdf.

[36] YAOLING NIU, MICHAEL J O'HARA and JULIAN A PEARCE. 'Initiation of subduction zones as a consequence of lateral compositional buoyancy contrast within the lithosphere: a petrological perspective'. In: *Journal of Petrology* 44.5 (2003), pp. 851–866.

[37] Cliff Ollier and Colin Pain. *The origin of mountains*. Routledge, 2004.

[38] Michael E Pasyanos. 'Lithospheric thickness modeled from long-period surface wave dispersion'. In: *Tectonophysics* 481.1 (2010), pp. 38–50.

[39] Eric T Peterson and Tetsuzo Seno. 'Factors affecting seismic moment release rates in subduction zones'. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 89.B12 (1984), pp. 10233–10248.

[40] *Plate Tectonics: Convergent plate boundaries & subduction-zone volcanism*. URL: http://www.volcanodiscovery.com/geology/subduction-zones.html.

[41] *Plate Tectonics Quiz*. URL: http://academic.brooklyn.cuny.edu/geology/grocha/plates/platequiz.htm.

[42] *Plates*. URL: http://www.ig.utexas.edu/research/projects/plates/.

[43] *Pplates Tectonic Reconstruction: The Mesh Paradigm*. 19th Apr. 2010. URL: http://rses.anu.edu.au/tectonics/projects/p-plates/deformable_mesh/index.php.

[44] Herbert Harold Read and Janet Watson. *INTRODUCTION TO GEOLOGY. VOLUME 1: PRINCIPLES*. Monograph. 1968.

[45] DH Roeder. 'Subduction and orogeny'. In: *Journal of Geophysical Research* 78.23 (1973), pp. 5005–5024.

[46] David B Rowley. 'Rate of plate creation and destruction: 180 Ma to present'. In: *Geological Society of America Bulletin* 114.8 (2002), pp. 927–933.

[47] Larry Ruff and Hiroo Kanamori. 'Seismicity and the subduction process'. In: *Physics of the Earth and Planetary Interiors* 23.3 (1980), pp. 240–252.

[48] WP Schellart and Nick Rawlinson. 'Global correlations between maximum magnitudes of subduction zone interface thrust earthquakes and physical parameters of subduction zones'. In: *Physics of the Earth and Planetary Interiors* 225 (2013), pp. 41–67.

[49] WP Schellart and Nick Rawlinson. 'Global correlations between maximum magnitudes of subduction zone interface thrust earthquakes and physical parameters of subduction zones'. In: *Physics of the Earth and Planetary Interiors* 225 (2013), pp. 41–67.

[50] Hans-Ulrich Schmincke. *Volcanism*. Vol. 28. Springer Science & Business Media, 2004.

[51] Maria Sdrolias and R Dietmar Müller. 'Controls on back-arc basin formation'. In: *Geochemistry, Geophysics, Geosystems* 7.4 (2006).

[52] H Sigurdsson et al. *Encyclopaedia of Volcanoes, 1417 pp*. 2000.

[53] Paul G Silver and Mark D Behn. 'Intermittent plate tectonics?' In: *science* 319.5859 (2008), pp. 85–88.

[54] *Snell's law*. 2015. URL: http://en.wikipedia.org/wiki/Snell's_law#mediaviewer/File:Snells_law2.svg.

[55] *SPlates Project*. URL: http://www.geodynamics.no/Web/Content/Projects/SPlates/.

[56] Seth Stein and Emile A Okal. 'Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process'. In: *Bulletin of the Seismological Society of America* 97.1A (2007), S279–S295.

[57] Seth Stein and Michael Wysession. *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons, 2009.

[58] Seth Stein and Michael Wysession. *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons, 2009.

[59] *Subduction*. URL: http://www.uoguelph.ca/geology/geol2250/glossary/HTML%20files/subduction.html.

[60] Toshiro Tanimoto and Thorne Lay. 'Mantle dynamics and seismic tomography'. In: *Proceedings of the National Academy of Sciences* 97.23 (2000), pp. 12409–12410.

[61] Wayne Thatcher. 'Silent slip on the Cascadia subduction interface'. In: *Science* 292.5521 (2001), pp. 1495–1496.

[62] *The layers of Earth*. 2003. URL: http://www.visionlearning.com/en/library/Earth-Science/6/Earth-Structure/69.

[63] *The Pplates Software Package*. 17th Mar. 2011. URL: http://rses.anu.edu.au/tectonics/projects/p-plates/.

[64] Aaron Tovish and Schubert Gerald. 'Island arc curvature, velocity of convergence and angle of subduction'. In: *Geophysical Research Letters* 5.5 (1978), pp. 329–332.

[65] *Trenches*. URL: http://www.whoi.edu/main/topic/trenches.

[66] DL Turcotte and ER Oxburgh. 'Finite amplitude convective cells and continental drift'. In: *Journal of Fluid Mechanics* 28.01 (1967), pp. 29–42.

[67] Seiya Uyeda et al. *new view of the Earth*. WH Freeman, 1978.

[68] Seiya Uyeda and Hiroo Kanamori. 'Back-arc opening and the mode of subduction'. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 84.B3 (1979), pp. 1049–1061.

[69] *Where and how do the continents deform?, Himalayan tectonics, Dynamic Earth*. Oct. 2001. URL: http://www.see.leeds.ac.uk/structure/dynamicearth/himalayas/deformation/index.htm.

[70] Simon E Williams et al. 'An open-source software environment for visualizing and refining plate tectonic reconstructions using high-resolution geological and geophysical data sets'. In: *GSA Today* 22.4/5 (2012), pp. 4–9.

[71] Takanobu Yokokura. 'On subduction dip angles'. In: *Tectonophysics* 77.1 (1981), pp. 63–77.